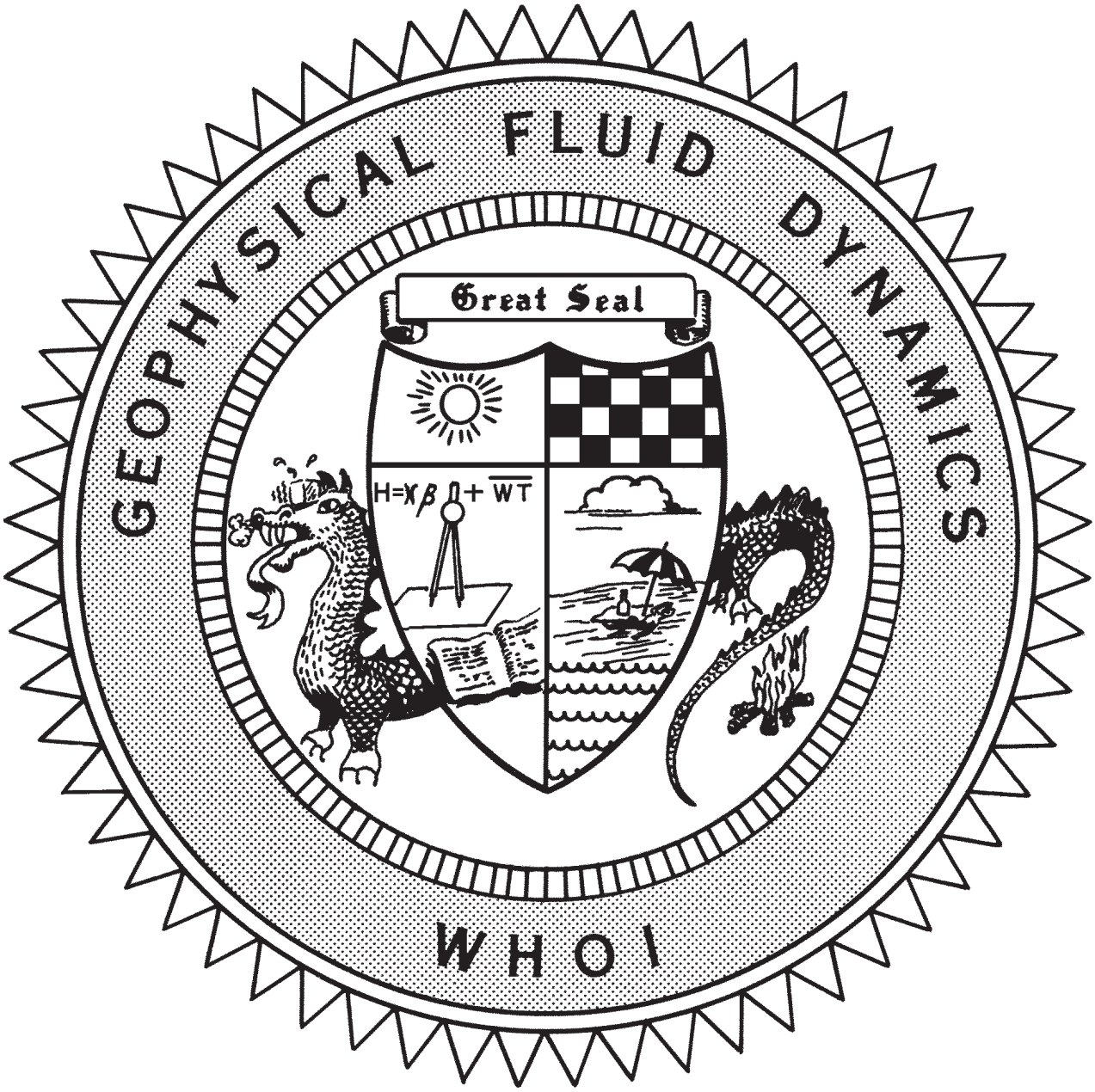
WHOI-2005-08

2004 Program of Study: Tides



Course Lectures
Fellows Project Reports

WHOI-2005-08

2004 Program of Study: Tides

by

Neil J. Balmforth and Stefan Llewellyn-Smith Co-Directors;
Myrl Hendershott and Christopher Garrett, Principal Lecturers

July 2005

Technical Report

Funding was provided by the Office of Naval Research under Contract No. N00014-04-1-0157 and the National Science Foundation under Grant No. OCE-0325296.

Approved for Distribution:

Kenneth W. Brink for Nelson Hogg

Nelson G. Hogg, Chair

Department of Physical Oceanography

Abstract

The summer of 2004 saw the GFD program tackle “Tides”. Myrl Hendershott (Scripps Institution of Oceanography) gave a fabulous introduction to the subject in the first week of the course, laying the foundations from astronomy and classical geophysical fluid dynamics. In the second week, Chris Garrett (University of Victoria) admirably followed up with recent developments on the subject, including the recent observations from satellite altimetry, their implications to mixing and circulation, and even a memorable lecture on the noble theme of how we might solve the world's energy crisis. The principal lectures proved unusually popular this summer, and the seminar room at Walsh often overflowed in the first two weeks.

Following on from the lectures, the seminar schedule of the summer covered in greater detail the oceanographic issues with which researchers are actively grappling. We also heard about related problems regarding atmospheric, planetary and stellar tides, together with the usual mix of topics on GFD in general.

The summer once again featured a lecture for the general public in the Woods Hole area. Carl Wunsch delivered a very well received lecture entitled “Climate Change Stories”, in which he gave an impression of how scientists generally believe our climate is currently changing, whilst simultaneously urging caution against some of the more outrageous and exaggerated claims. The lecture was held at Lilly Auditorium, thanks to the hospitality of the Marine Biology Laboratory. The reception following the lecture was enjoyed by all.

Neil Balmforth and Stefan Llewellyn Smith acted as Co-Directors for the summer. Janet Fields, Jeanne Fleming and Penny Foster provided the administrative backbone to the Program, both during the summer and throughout the year beforehand. As always, we were grateful to the Woods Hole Oceanographic Institution for the use of Walsh Cottage, and Keith Bradley's solid service could not be overlooked. Shilpa Ghadge and Shreyas Mandre are to be thanked for their part in comforting the fellows, developing the summer's proceedings volume (available on the GFD web site) and for running the computer network.

TABLE OF CONTENTS

2004 GFD FELLOWS, STAFF AND VISITORS

Fellows

Josefina Arraut	Brazilian Institute for Space Research
Vineet Birman	University of California, Santa Barbara
Eleanor Williams Frajka	University of Washington
Visweswaran Nageswaran	University of Massachusetts, Amherst
Lisa Neef	University of Toronto
Anja Slim	University of Cambridge
Yaron Toledo	Technion, Israel
David Vener	Massachusetts Institute of Technology
Danielle Wain	University of Illinois
Marshall Ward	Florida State University

Staff and Visitors

Brian Arbic	Princeton University
Phil Arras	University of California, Santa Barbara
Neil J. Balmforth	University of British Columbia
Andrew Belmonte	Penn State University
Andrew Bernoff	Harvey Mudd College
Gregory Buck	Saint Anselm College
John Bush	Massachusetts Institute of Technology
Stephen W. Childress	Courant Institute, New York University
Colin Cotter	Imperial College
Richard Craster	Imperial College
Charles Doering	University of Michigan
Gary Egbert	Oregon State University
Raffaele Ferrari	Massachusetts Institute of Technology
Glen Flierl	Massachusetts Institute of Technology
Jeffrey Forbes	University of Colorado, Boulder
Christopher Garrett	University of Victoria
Shilpa Ghadge	University of British Columbia
Sara Gille	University of California, San Diego
Harvey Greenspan	Massachusetts Institute of Technology
Pin-Gao Gu	Academia Sinica
Maura Hagan	NCAR
Myrl Hendershott	University of California, San Diego
Annette Hosoi	Massachusetts Institute of Technology
Louis N. Howard	Florida State Univ./Massachusetts Institute of Technology
Herbert Huppert	University of Cambridge
Joseph B. Keller	Stanford University
Richard Kerswell	University of Bristol
Jody Klymak	University of California, San Diego
Eric Kunze	University of Washington
Xinfeng Lang	Ocean University of China
Norman R. Lebovitz	University of Chicago
Zhi Li	Ocean University of China

Doug Lin	University of California
Jennifer A. MacKinnon	University of California, San Diego
Amala Mahadevan	Boston University
Lakshminarayanan Mahadevan	Harvard University
Willem V.R. Malkus	Massachusetts Institute of Technology
John Marshall	Massachusetts Institute of Technology
Shreyas Mandre	University of British Columbia
Leo Maas	NIOZ
Mark Merrifield	University of Hawaii
Philip J. Morrison	University of Texas at Austin
Thomas Mullin	University of Manchester
Jonas Nycander	Stockholm University
Gordon Ogilvie	Institute of Astronomy
Richard Pawlowicz	University of British Columbia
Thomas Peacock	Massachusetts Institute of Technology
Robert Pinkel	University of California, San Diego
Francis Poulin	Scripps Institution of Oceanography
Luc Rainville	Scripps Institution of Oceanography
Alan Rempel	Massachusetts Institute of Technology
Claes G. Rooth	University of Miami
Harper Simmons	University of Alaska, Fairbanks
Stefan Llewellyn-Smith	University of California, San Diego
Alexander Soloviev	NOVA Southeastern University
Tilman Spohn	Westfaelische Wilhelms Universitaet
Edward A. Spiegel	Columbia University
Roman Stocker	Massachusetts Institute of Technology
Bruce Sutherland	University of Alberta
Jean-Luc Thiffeault	Imperial College
Eli Tziperman	Weizmann Institute of Science
George Veronis	Yale University
John Wahr	University of Colorado, Boulder
Matthew Wells	Yale University
John Wettlaufer	Yale University
Kraig Winters	University of California, San Diego
Carl Wunsch	Massachusetts Institute of Technology

2004 Participants

Standing (left to right): P. Foster, S. Ghadge, S. Llewellyn-Smith, L. Howard, A. Belmonte, N. Lebovitz, P. Cessi, J. MacKinnon, M. Ward, V. Birman, A. Provenzale, C. Cenedese, G. Buck, G. Oglivie, S. Jayne

Seated on porch (left to right): A. Slim, G. Egbert, S. Gille, E. Williams Frajka, J. Arraut, G. Veronis, D. Wain, D. Vener, L. Neef, Y. Toledo, V. Nageswaran, W. Young

Seated on ground (left to right): P. Gu, J-L. Thiffeault, J. Keller, S. Mandre, N. Balmforth

2004 Principal Lecturer Myrl Hendershott

2004 Principal Lecturer Christopher Garrett

GFD 2004 Lecture Schedule

(All talks held in Walsh cottage unless otherwise noted.)

Week of June 21 – 25, 2004

PRINCIPAL LECTURES Myrl Hendershott, Scripps Institution of Oceanography

Monday, June 21
10:00 am ***Introduction to Ocean Tides***

Tuesday, June 22
10:00 am ***The Role of Tidal Dissipation and the Laplace Tidal Equations***

Wednesday, June 23
10:00 am ***Solutions to Laplace's Tidal Equations***

Thursday, June 24
10:00 am ***Resonance and Solutions to the LTE***

Friday, June 25
10:00 am ***The Spectrum of Free Waves Possible along Coasts***

Week of June 28 – July 2, 2004

PRINCIPAL LECTURES Christopher Garrett, University of Victoria, Canada

Monday, June 28
10:00 am ***Internal Tides***

Tuesday, June 29
10:00 am ***Tidal Bores***

Wednesday, June 30
10:00 am ***Tidal Rectification and Stokes Drift***

Thursday, July 1
10:00 am ***Tidal Rectification, Stratification and Mixing***

Friday, July 2
10:00 am ***Tidal Power***

Week of July 5 – 9, 2004

Monday, July 5

NO LECTURE – DUE TO INDEPENDENCE DAY HOLIDAY

Tuesday, July 6

10:30 am Jean-Paul Zahn, Observatoire de Paris
Tidal Evolution of Binary Stars

Wednesday, July 7

10:30 am Jonas Nycander, Stockholm University
Generation of Internal Waves in the Deep Ocean by Tides

Thursday, July 8

10:30 am Sarah Gille and Stefan Llewellyn-Smith, University of California
Global Land and Sea Breeze

2:30 pm Jennifer Mackinnon, Scripps Institution of Oceanography
A Primer on Internal Waves

Friday, July 9

10:30 am Ed Spiegel, Columbia University
On the Bifurcation of Species

Week of July 12 – 16, 2004

Monday, July 12

10:30 am Rob Pinkel, Scripps Institution of Oceanography
Home

2:30 pm Luc Rainville, Scripps Institution of Oceanography
Internal Wave Propagation from the Hawaiian Ridge

Tuesday, July 13

10:30 am Maura Hagan, National Center for Atmospheric Research
Thermal Tides in Earth's Atmosphere

2:30 pm Jeff Forbes, University of Colorado
Nonmigrating Tides in Planetary Atmospheres

Wednesday, July 14

10:00 am Carl Wunsch, Massachusetts Institute of Technology
Ocean Tides and Climate Change

Thursday, July 15

10:30 am Brian Arbic, Princeton University
Global Barotropic and Baroclinic Tide Models

3:30 pm Peter Rhines, Steinbach Scholar – P.O. Seminar – Carriage House
Rossby Waves and Potential Vorticity in Oceans and Atmospheres

Friday, July 16

10:30 am James Girton, Physical Oceanography Department, W.H.O.I.
Internal Tide Generation and Dissipation: A Nearfield View

3:00 pm Peter Rhines, Steinbach Scholar – Institute Seminar – Clark 507
Exploring the Subpolar Atlantic: Climate, Dynamics and Seagliders

Week of July 19 – 23, 2004

Monday, July 19

10:30 am Jean-Luc Thiffeault, Imperial College, London
Knitting and Mixing

Tuesday, July 20

10:30 am Doug Lin, University of California
Tidal Interaction between Extrasolar Planets and their Host Stars

3:00 pm Harper Simmons, IARC, University of Alaska Fairbanks
P.O. Department Seminar, Clark 507
Baroclinic Wave Generation in a Tidal Model

Wednesday, July 21

10:30 am Bruce Sutherland, University of Alberta
Internal Wave Propagation: Ray Theory and Beyond

2:30 pm Jody Klymak, Scripps Institution of Oceanography
The Tidal Graveyard: Turbulent Dissipation near the Hawaiian Ridge

Thursday, July 22

10:30 am Richard Kerswell, University of Bristol, United Kingdom
Io, Tides and Magnetic Fields

3:00 pm Anand Gnandesikan, GFDL, Princeton University
Special P.O. Department Seminar, Clark 507
Connecting the ACC and THC: The Role of Southern Ocean Winds and Eddies

Friday, July 23

10:30 am Tom Mullin
Sticky Balls

2:30 pm Colin Cutter, Imperial College, London
The Imperial College Ocean Model: Simulating the Oceans using Finite Elements

Week of July 26 – 30, 2004

Monday, July 26
10:30 am Sonya Legg, Physical Oceanography Department, WHOI
Internal Tide Reflection from Topography

2:30 pm Steve Childress, Courant Institute
Geometric Constraints on Vortex Stretching

Tuesday, July 27
10:30 am Alexey Fedorov, Princeton University
Generation of Breaking Waves by Strong Winds

Wednesday, July 28
10:30 am Gary Egbert, Oregon State University
Long-period Tides

2:30 pm Gordon Ogilivie, University of Cambridge
Tidal Dissipation in Rotating Giant Planets

Thursday, July 29
10:30 am Andrew Belmonte, Pennsylvania State University
Viscoelastic Craters

Friday, July 30
10:30 am Pin-Gao Gu, Academia Sinica, China
Thermal Adjustment of Hot Jupiters due to Tidal Dissipation

Week of August 2 – 6, 2004

Monday, August 2
10:30 am Joel Miller, University of Cambridge, United Kingdom
Zero Reynolds Number Elastic Fluid Instabilities

2:30 pm Antonello Provenzale, ICG Torino, Italy
Two Short Stories in Turbulent Convection: The Clustering of Plumes and a Model of Atmospheric Convection

Tuesday, August 3

10:30 am Rui Xin Huang, Physical Oceanographic Department, WHOI
Energetics of the Oceanic Circulation: Sandstrom Theory and Lab Experiments

Wednesday, August 4

10:30 am Carl Wunsch, Massachusetts Institute of Technology
Climate Change Stories - A public lecture at Lily Auditorium, Reception to follow

Thursday, August 5

10:30 am Amala Mahadevan, Boston University
Phytoplankton, Nutrients and CO2 in the Surface Ocean

2:30 pm L. Mahadevan, Harvard University
Particulate Interfaces

Friday, August 6

10:30 am Charles Doering, University of Michigan
A Bound on Mixing Efficiency

Week of August 9 – 13, 2004

Monday, August 9

10:30 am Leo Maas, NIOZ, The Netherlands
Wave Attractors

Tuesday, August 10

10:30 am Matthew Wells, Yale University
Laboratory Experiments on Entrainment in Gravity Currents with Stratification and Rotation

2:30 pm Jennifer MacKinnon, Scripps Institution of Oceanography
Physical Oceanography Department Seminar, Clark 507
Mixing Hotspots and other Catastrophes: Mechanisms of Tidal Dissipation

Wednesday, August 11

10:30 am Norman Lebovitz, University of Chicago
Elliptical Instabilities

Thursday, August 12

2:30 pm Paul Dellar, Imperial College, London
The Revolutionary Shallow Water Equations

Friday, August 13 **NO LECTURE**

Week of August 16 – 20, 2004 - Quiet Week, No Lectures

Week of August 23 – 27, 2004 - Fellows Lectures

Tuesday, August 24

10:00 am Danielle Wain, University of Illinois
Laboratory Experiments on the Effects of Baroclinic Eddies on a Dense Plume in a Rotating Stratified Field

11:15 am Eleanor Williams Frajka, University of Washington
Convection in a Fluid Loop

2:00 pm Lisa Neef, University of Toronto
Resonant Triads of Tidally-Forced Internal Gravity Waves

3:15 pm Yaron Toledo, Technion, Israel
High Order Boussinesq Models for Internal Interfacial waves and Layered SQG

Wednesday, August 25

10:00 am Anja Slim, University of Cambridge, United Kingdom
Exotic Gravity Currents

11:15 am Josefina Arraut, Brazilian Institute for Space Research
Rossby Wave Scattering between Homogeneous Media

2:00 pm Marshall Ward, Florida State University
Double Diffusive Gravity Currents under Rotation

3:15 pm David Vener, Massachusetts Institute of Technology
Two-Dimensional Vortex Shedding from a Corner

Thursday, August 26

10:00 am Vineet Berman, University of California, Santa Barbara
Non-Normal Convection

11:15 am Visweswaran Nagewswaran, University of Massachusetts, Amherst
Parametric Instability of Internal Waves with Rotation

Lecture 1: Introduction to ocean tides

Myrl Hendershott

1 Introduction

The phenomenon of oceanic tides has been observed and studied by humanity for centuries. Success in localized tidal prediction and in the general understanding of tidal propagation in ocean basins led to the belief that this was a well understood phenomenon and no longer of interest for scientific investigation. However, recent decades have seen a renewal of interest for this subject by the scientific community. The goal is now to understand the dissipation of tidal energy in the ocean. Research done in the seventies suggested that rather than being mostly dissipated on continental shelves and shallow seas, tidal energy could excite far traveling internal waves in the ocean. Through interaction with oceanic currents, topographic features or with other waves, these could transfer energy to smaller scales and contribute to oceanic mixing. This has been suggested as a possible driving mechanism for the thermohaline circulation.

This first lecture is introductory and its aim is to review the tidal generating mechanisms and to arrive at a mathematical expression for the tide generating potential.

2 Tide Generating Forces

Tidal oscillations are the response of the ocean and the Earth to the gravitational pull of celestial bodies other than the Earth. Because of their movement relative to the Earth, this gravitational pull changes in time, and because of the finite size of the Earth, it also varies in space over its surface. Fortunately for local tidal prediction, the temporal response of the ocean is very linear, allowing tidal records to be interpreted as the superposition of periodic components with frequencies associated with the movements of the celestial bodies exerting the force. Spatial response is influenced by the presence of continents and bottom topography, and is a less well established matter.

Figure 1 shows a two month tidal record from Port Adelaide, Australia. Even though tidal records vary significantly for different coastal locations, this one in particular can be considered typical in that it clearly shows characteristics of tidal oscillations that can be directly related to astronomical forcings.

Perhaps the first feature to stand out is the semi-diurnal component, two high tides can be seen to occur on each day. A closer look reveals a modulation of the amplitude of the semi-diurnal oscillation, roughly over a one month period. Intervals of high amplitude are known as spring tides while those of lower amplitudes are known as neap tides. As indicated in the figure, the springs-neaps cycle is associated with the phases of the Moon. For a same day, there is often a difference in the amplitude of the two high tides. This is known as the

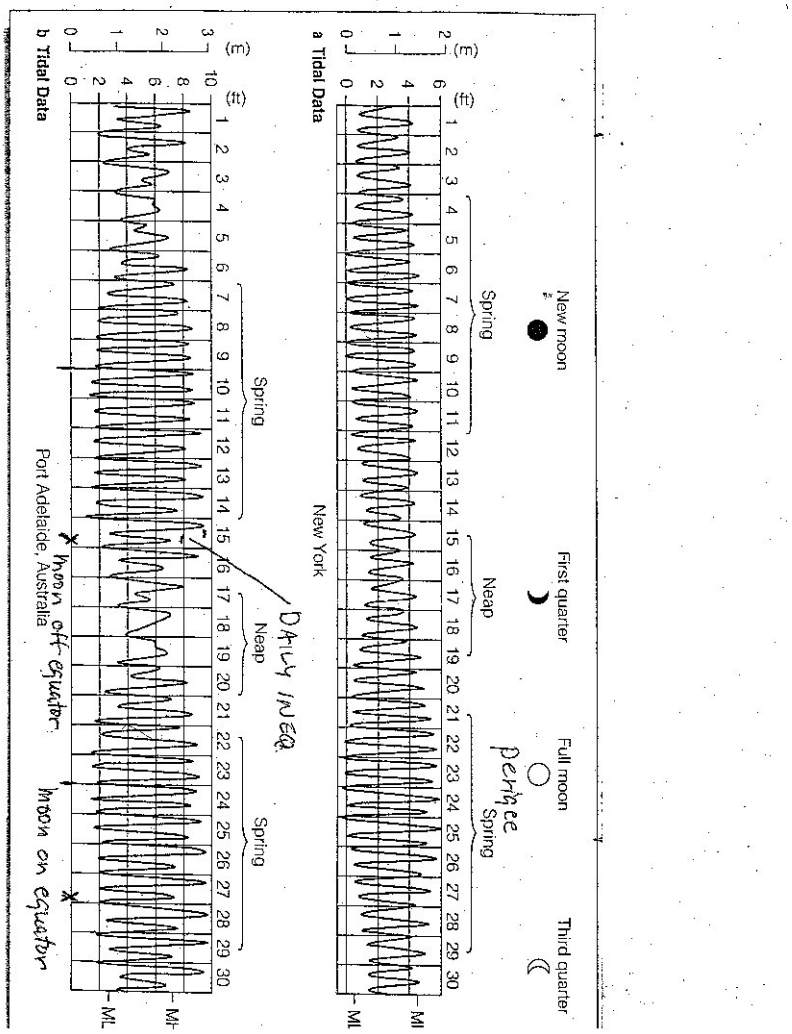


Figure 1: Two months of tidal data for Port Adelaide Australia. We can see in this record some features that can be directly accounted for by the details of the astronomical tidal forcing, such as the springs-neaps cycle, the daily inequality and the absence of daily inequality when the Moon is on the Equator.

"daily inequality" and, as indicated in the figure, disappears when the Moon is over the equator.

We will try to clarify below the basic ideas underlying tidal forcing. We will also try to explain the origin of the forcing terms responsible for causing the tidal record features described above. Finally we will attempt to explain the derivation of the tide generating potential.

3 Tidal Forcing

Even though small compared to the planets and especially to the sun, the Moon is by far the celestial body closest to the Earth. Because gravitational pull decreases linearly with mass and quadratically with distance, the Moon exerts the biggest influence over the Earth, contributing the most to the formation of tides. We will begin by considering its effects.

The centres of mass of the Earth and the Moon orbit around the common centre of mass of the Earth-Moon system. Their movements are such that centrifugal force counterbalances gravitational attraction at the individual centres of mass. The Earth is a rigid body, so every material point in it executes an identical orbit, and is therefore subject to the same centrifugal force, as illustrated in figure 2. Gravitational force however will vary because the distance between these points to the Moon may vary by up to one Earth diameter. Gravitational force will prevail over centrifugal force on the hemisphere closest to the Moon and centrifugal force will prevail on the hemisphere furthest to it. The opposite hemispheres have net forces in opposite directions, causing the ocean to bulge on both sides. As the Earth spins under this configuration, two daily tides are felt.

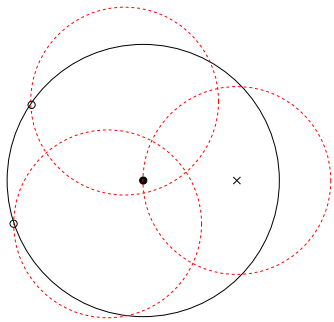


Figure 2: The centre of the Earth, shown as a filled dot, rotates about the centre of mass of the Earth-Moon system, indicated by a 'x' mark. Dashed circles show the orbital movement of the points shown by a 'o' mark on Earth's surface and the center of the Earth.

A common source of confusion regarding the argument above is to suppose that the centrifugal force relevant to the problem is due to the spinning of the Earth around its own axis. This seems reasonable at first sight because this force is constant for every latitude circle, allowing for an imbalance with lunar attraction, which is longitude dependent at any given instant. However, the centrifugal force due to the Earth's spin has permanently deformed the Earth's surface into a spheroid (as opposed to the spherical shape that would ensue from self-gravitation only). That is to say, this centrifugal force is compensated by

the Earth's own gravitational field. A mnemonic phrase to keep in mind is that tides are caused by the action of *other* celestial bodies over the Earth.

Another not entirely uncommon misconception is that tides are partly the result of the variation of centrifugal force over the surface of the Earth. This kind of confusion arises because the distance between the centre of the Earth-Moon system is smaller than one Earth radius and therefore this point lies "within" the Earth. If this were a fixed material point, like the centre of the Earth, around which the planet revolved, there would indeed be a variation of centrifugal force with distance from it. However, the centre of mass of the Earth-Moon system is just a point in space, and as the Earth revolves around it, as indicated in figure 2 none of its material points are fixed.

Even though a constant field, the centrifugal force due to the revolution of the Earth around the system's centre of mass is essential to the semi-diurnality of the tides. We can illustrate this by considering the situation in which the centre of the Earth is fixed. In this case the only force acting upon it is lunar gravitational attraction. Although uneven over the surface, it pulls every point on Earth towards the Moon, causing the water to bulge on the hemisphere closer to it. The spinning of the Earth would therefore make every point on it experience a diurnal tide cycle, instead of a semi-diurnal one. This situation is illustrated in figure 3

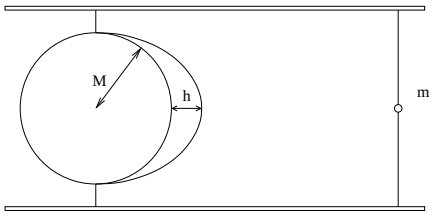


Figure 3: Tidal deformation that would ensue if from lunar attraction if the Earth's centre of mass were fixed.

We can calculate the surface elevation that would result from this forcing. The assumption is that the elevation would be such that net terrestrial gravitational force would exactly compensate for the lunar force at the point.

Net Moon's gravitational force:

$$\frac{GM}{(a+h)^2} - \frac{GM}{a^2} \simeq -\frac{GM}{a^2}\frac{2h}{a}.$$

Net Earth's gravitational force:

$$\frac{GM}{a^2}(\frac{a}{r})^2\frac{m}{M},$$

where M is the mass of the Earth, r is the distance between Earth's and Moon's centres of mass, a is the Earth's radius, h is the sea surface deformation at the sub lunar point and G is the universal gravitational constant. Equating the two forces and isolating h we get:

$$h = \frac{a^3}{2r^2}\frac{m}{M} = 10.7\text{m}. \tag{1}$$

This is an unrealistically high value.

If we now allow the Earth's centre of mass to accelerate, the centrifugal force due to this motion will compensate for the Moon's gravity at that point. We can anticipate that tidal deformation will be smaller, since it will be a response to a smaller resultant force. As explained earlier, predominance of lunar attraction on the hemisphere facing the Moon and of the centrifugal force on the one opposing deforms the surface into an ellipsoid. The water will bulge around the sub-lunar and anti-sub lunar points.

For this case we can also calculate tidal elevation on the sub-lunar point.

$$\frac{Gm}{(r-a)^2} - \frac{Gm}{r^2} \simeq \frac{GM}{r^2}\frac{2a}{r}, \tag{2}$$

where the first term on the right is lunar gravity, the second is the centrifugal force and the term on the right is the net gravitational force of the Earth. Isolating h we get:

$$h = \frac{a^4}{r^3}\frac{m}{M} = 35.8\text{cm} \tag{3}$$

The Moon rotates around the Earth in the same direction as the Earth spins, and the surface deformation must rotate with it. It takes slightly longer than a day for the Moon to be directly over the same point on the Earth's surface, as illustrated in figure 4 this is called a lunar day. Likewise, the period between two high tides is half lunar day. Apart from the semi-diurnal tides, we can expect the presence of the Moon to permanently deform the sea surface. This is an order zero effect called the permanent tide.

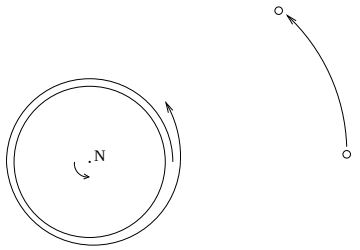


Figure 4: The Earth-Moon system. The position of the Moon with respect to a fixed point on the Earth's surface after one revolution is illustrated.

Up to now, we have considered the orbit of the Moon circular. It is however elliptic, with the Earth-Moon centre of mass being one of the foci. The Moon's gravitational force over the Earth will be modulated over the period of one anomalistic month (figure 5), as will the tidal components it generates.

In summary, tidal forcing by the Moon alone can be represented by the following harmonics:

Lunar Semi-Diurnal Tide (M_2) $2/LD$	12h 25.236 min
Lunar Elliptical (N_2)	$2/LD - 1/perigee$ 12h 3.501 min
Lunar Monthly Elliptical (M_m)	$1/perigee$ 27.5545 days (anomalistic month)

The $2/LD + 1/perigee$ term was left out because it has a small amplitude. Modulation of the amplitude of M_2 is represented by the interaction of M_2 and N_2, which is constructive once each anomalistic month.

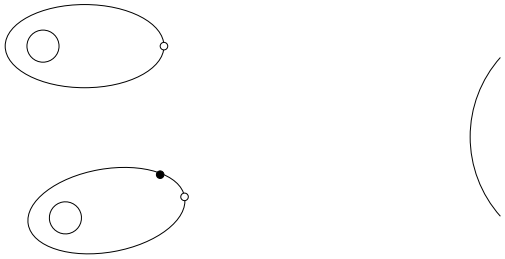


Figure 5: As the Moon rotates around the Earth, the Earth rotates around the Sun. For this reason, after each orbital period (anomalistic month), the Moon is in different position in its orbit with respect to the Sun. In the above picture, we represent on top an initial position of the Earth and Moon, below it their position after one anomalistic month. The black Moon in this case represents the position it would have to be in to exhibit the same phase as in the initial configuration.

All arguments mentioned above are valid for any other celestial body which might be reasonably considered to form a two body system with the Earth, for which the orbits are elliptical. Another such body is the sun, whose tidal effect over the Earth (when the two body system is considered in isolation) can be reduced to the harmonic components below.

Solar Semi-Diurnal Tide (S_2)	$2/SD$ 12h 25.236 min
Solar Elliptical (N_2)	$2/SD - 1/anom.yr.$ 12h 3.501 min
Solar Annual Elliptical ()	$1/perihelion$ 365.25964 days (anomalistic year)

When Moon and Sun are aligned with the Earth, their semi-diurnal components interfere constructively, giving rise to tides of larger amplitude, known as spring tides. When they are in quadrature, the interference is exactly destructive, giving rise to smaller amplitude tidal variations, called neap tides. The relative arrangement of the Earth Sun and Moon is perceived on the Earth as the phases of the Moon, and therefore the springs-neaps cycle has a period of one lunar (synodic) month. A lunar month is the duration required for the Moon to return to a fixed position in its orbit in relation to the Sun as illustrated in figure 5.

Up to now we have assumed that the orbits of the Earth (around the Sun) and Moon are coplanar to the spinning of the Earth at all instants of time. In reality these planes intersect at an angle. The effect this has over the tide is illustrated in figure 6. As the Earth rotates,

it perceives the tidal surface as being "tilted" in relation to latitude. In terms of harmonics, this is represented by a daily component, which gives rise to the daily inequality. When the tide generating bodies intersect the equatorial plane, the daily inequality disappears. In the Port Adelaide record (figure 1), we can see that daily inequality disappears when the Moon is on the Equator.

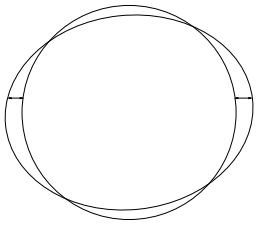


Figure 6: The Moon's orbit is not coplanar with the Equator. The tidal surface is in general "tilted" with respect to the Equator giving rise to the daily inequality.

The amplitude of the declinational components depends on the angle of the bodies orbit to the equatorial plane. Both the lunar orbital plane and the ecliptic precess, modulating the declinational tides. The Moon's orbit precesses over 18.6 years, its angle to the Ecliptic varying between $-5°08'$ and $5°08'$. In relation to the Earth's equatorial plane the variation is between $23°27' - 5°08'$ and $23°27' + 5°08'$.

As mentioned earlier, the usefulness of decomposing the tide generating force into harmonics is due to the linearity of the oceans response to it in time. In fact, we have taken this for granted in the preceding section when we explained the springs-neaps cycle purely as the result of the interference of two forcing terms. This property allows for more precise tidal prediction. Tidal records are not used to determine the important frequencies in their harmonic expansions, these are known from astronomical considerations. Data is used only to determine the amplitudes of local response to these terms.

4 Spatial Structure of the Tides

As the Earth's spinning under the tide generating potential is felt as the propagation of the tidal wave. However, this propagation is obstructed by the presence of continents and bottom topography. Real co-tidal lines therefore look nothing like the constant phase lines of the tide generating potential. As a plane wave enters a basin, it feels the effect of the Earth's rotation and propagates along its borders. The nodal line that would exist in the case with no rotation degenerates into a nodal point, called amphidromic point.The irregularity of the oceanic basins and of bottom topography disrupt the propagation and a precise map of tidal propagation could only be obtained after the advent of satellite altimetry. This data is harmonically analyzed to obtain maps for the different astronomical components. Figure 7 shows a co-tidal map obtained in this manner. Although the amphidromic points are eye catching, the less conspicuous anti-amphidromic points, for which tidal amplitude is maximum and there is almost no phase variation, are probably more useful for testing satellite altimetry.

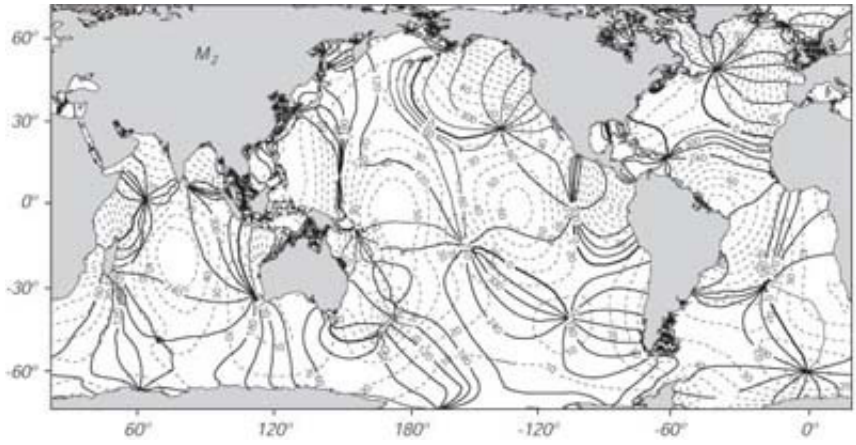


Figure 7: Cotidal map. Amphidromic points are the ones where cotidal lines cross. In these points tidal amplitude is zero and phase speed infinite. Less conspicuous are the anti-amphidromic points, where tidal amplitude is maximum and phase stationary.

In the following section we will formalize the ideas outlined above so as to arrive at an expression for the tide generating potential.

5 The tide generating potential

We want to calculate the tidal force that the Moon or Sun exerts on the Earth, in particular on the oceans. Remember that we are only interested in the effects of *another* body on the Earth, not the effect of the Earth's rotation and gravity on its shape and that of the oceans.

Consider the plane made up of the centre of the Earth, the tide generating body (the Moon or the Sun), assumed to be a point mass, and an observer at P on the surface of the Earth, as shown in figure 8. For the moment, we assume R is constant.

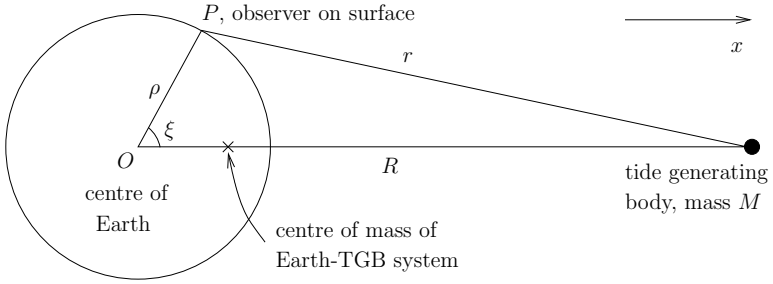


Figure 8: The geometry for calculating the tidal force at P due to the tide generating body.

The tide generating body (TGB) exerts a force, $\mathbf{F}$, on the centre of the Earth

$$\mathbf{F} = \frac{GM}{R^2}\,\hat{\imath},$$

where $\hat{\imath}$ is the unit normal along the x-direction and G is the universal gravitational constant. From this, the potential at the centre of the Earth, O, due to the TGB is

$$V(O) = \frac{GM}{R^2}\, x + \text{const} = \frac{GM}{R^2}\rho\cos\xi + \text{const},$$

where the constant is arbitrary.

Then the resultant force at P is the gravitational force at P, less that at O, with potential

$$V(P) = \frac{GM}{r} - \frac{GM}{R^2}\rho\cos\xi. \tag{4}$$

The sign is chosen as r is measured away from the TGB and the force should be towards it. Note that the force on the Earth as a whole, $\mathbf{F}$, is balanced by the centrifugal force due to the motion of the Earth about the common centre of mass of the Earth–TGB system.

The parameters r, R, ρ and ξ are linked by $r^2 = R^2 - 2\rho R\cos\xi + \rho^2$ from properties of triangles. Hence

$$\frac{1}{r} = \frac{1}{R}\left(1 - \frac{2\rho}{R}\cos\xi + \frac{\rho^2}{R^2}\right)^{-\frac{1}{2}} = \frac{1}{R}\sum_{n=0}^{\infty}\left(\frac{\rho}{R}\right)^n P_n(\cos\xi), \tag{5}$$

where $P_n(z)$ is the nth Legendre polynomial, with $P_0(z) = 1$, $P_1(z) = z$, $P_2(z) = (3z^2-1)/2$, ... The final equality in (5) may be found in, for example, Morse and Feshbach [1].

Hence (4) becomes

$$V(P) = \frac{GM}{R}\left[1 + \sum_{n=2}^{\infty}\left(\frac{\rho}{R}\right)^n P_n(\cos\xi)\right].$$

For the Moon, $0.0157 \leq \rho/R \leq 0.0180$ and for the Sun $\rho/R \sim 10^{-4}$. Hence the potential may be truncated at $n = 2$. Forgetting about the constant GM/R, which is unimportant since ultimately we want to find the forces, it becomes

$$V(P) \simeq \frac{GM}{R}\left(\frac{\rho}{R}\right)^2 P_2(\cos\xi). \tag{6}$$

Note that this potential is symmetric in ξ. This is consistent with the discussion in previous sections, where we argued that the tide generating force is symmetrical with respect to the plane that contains the Earth's centre of mass and is orthogonal to the Earth-Moon axis.

5.1 The tide generating potential in geographical coordinates

It is more useful to express the tidal potential in geographical coordinates: actual latitude and longitude of the observer on the Earth and the apparent latitude and longitude of the TGB. This coordinate system is shown in figure 9. We call attention to the fact that the longitudinal angles are measures with respect to the Equator. In this coordinate system the tidal ellipsoid is "tilted", and the tidal potential will therefore have asymmetrical components.

From spherical trigonometry

$$\cos\xi = \sin\theta\sin\delta + \cos\theta\cos\delta\cos H.$$

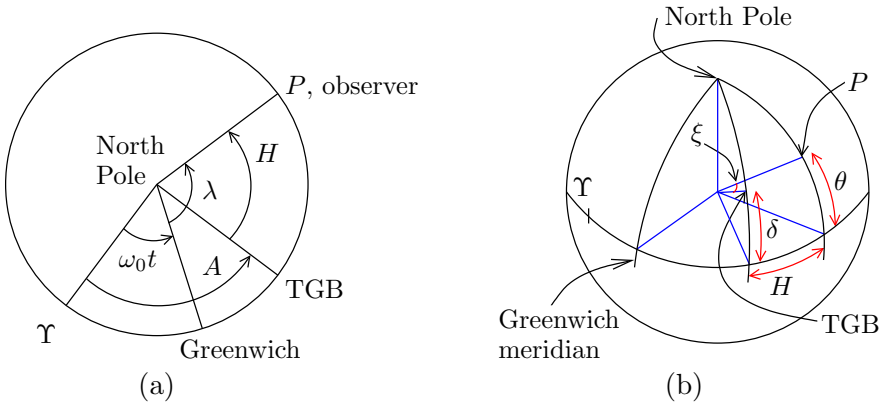


Figure 9: The geographical coordinates: (a) looking down from the North Pole and (b) looking at the spherical Earth. Here λ is the geographic longitude of the observer and θ is the geographic latitude. ξ is the angle between the TGB and the observer as in figure 8. δ is the *declination* of the TGB and A is the *right ascension* of the TGB, or its apparent longitude with respect to an origin Υ. Υ is a celestial reference point from which to measure positions of the TGB: it is the northward crossing of the Sun at equinox, or equivalently the point of intersection of the equatorial plane and the *ecliptic*, the plane of the Earth's orbit about the Sun, as the Sun travels northward. The point Υ is assumed fixed with respect to the fixed stars for the current discussion. However, with respect to the rotating Earth, the point Υ moves. If $t = 0$ is taken to be the time at which Υ lies on the Greenwich meridian, then at time t, the Earth has rotated $\omega_0 t$ giving the angle between Υ and Greenwich, where ω_0 is the frequency associated with the period of rotation of the Earth about its axis so that the TGB appears again in the same Earth–TGB orientation. $H = \omega_0 t + \lambda - A$ is referred to as the hour angle.

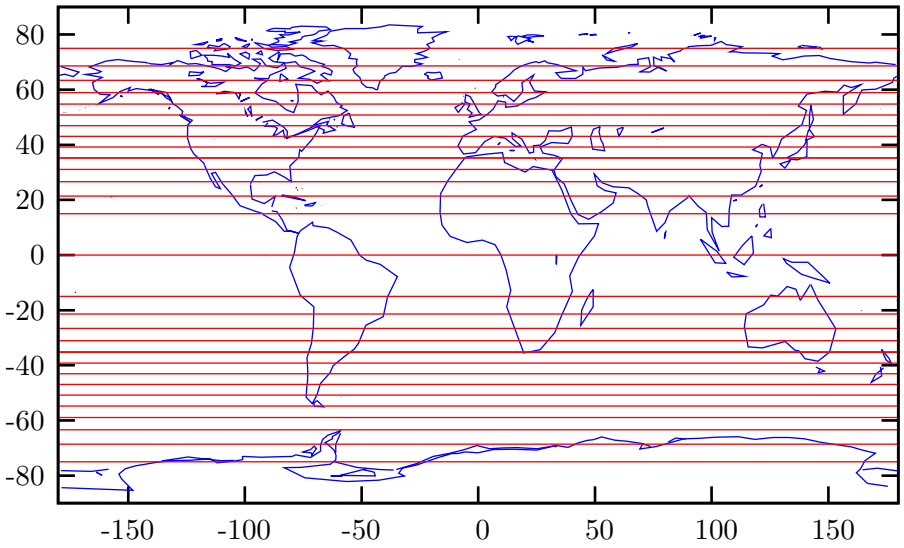


Figure 10: Equipotential lines for the long-period potential.

Substituting this in (6) gives

$$V(\lambda,\theta) \simeq \frac{3GM\rho^2}{4R_0^3}\left(\frac{R_0}{R}\right)^3\left[\frac{4}{3}\left(\frac{1}{2}-\frac{3}{2}\sin^2\theta\right)\left(\frac{1}{2}-\frac{3}{2}\sin^2\delta\right)\right.$$
$$\left.+\sin 2\theta\sin 2\delta\cos H+\cos^2\theta\cos^2\delta\cos 2H\right], \qquad (7)$$

where R_0 is a reference value of the orbital distance R of the TGB. The coefficient $3GM\rho^2/4R_0^3$ is referred to as the *Doodson constant*, D. For the Moon, $D_{\rm Moon}/g = 26.75\,$cm and for the Sun, $D_{\rm sun} = 0.4605 D_{\rm Moon}$.

The first term in the square bracket in (7) has no dependence on the hour angle, H, and gives rise to a long-period[1] potential. The second term gives rise to a diurnal potential and the final term to a semi-diurnal potential. Figures 10-12 show plots of the instantaneous equipotential lines for the three components of (7) and plots of the cotidal[2] and corange[3] lines. The plots shown are in fact representative of the solid Earth tide, since the response time of the Earth is of the order of an hour (deduced from earthquake measurements), much quicker than the time period of the tidal potential and so the solid Earth can adjust to the equipotential surfaces. The oceans have a much longer response time.

In reality, the declination and orbital distance in (7) vary in time. In the following sections, we consider how these variations change the potential.

[1] In the present analysis, infinitely long. However we shall see that each term of the potential is modulated and hence this becomes a long-period potential.

[2] A line passing through points at which high tide occurs at the same number of hours after the Moon transits the Greenwich meridian.

[3] A line passing through points of equal tidal range.

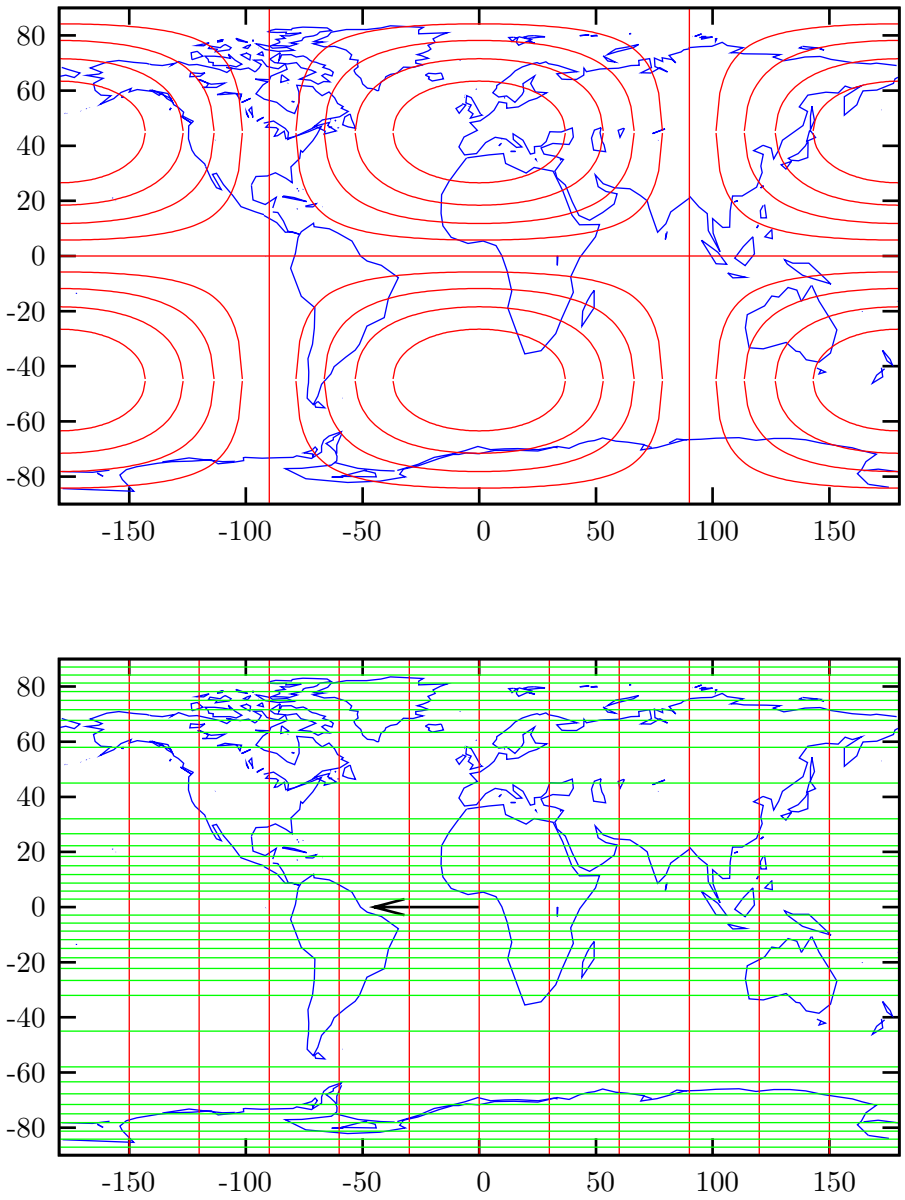


Figure 11: Equipotential lines (top) and cotidal, green, and corange, red, lines (bottom) for the diurnal potential.The arrow indicates the direction of increasing cotidal time.

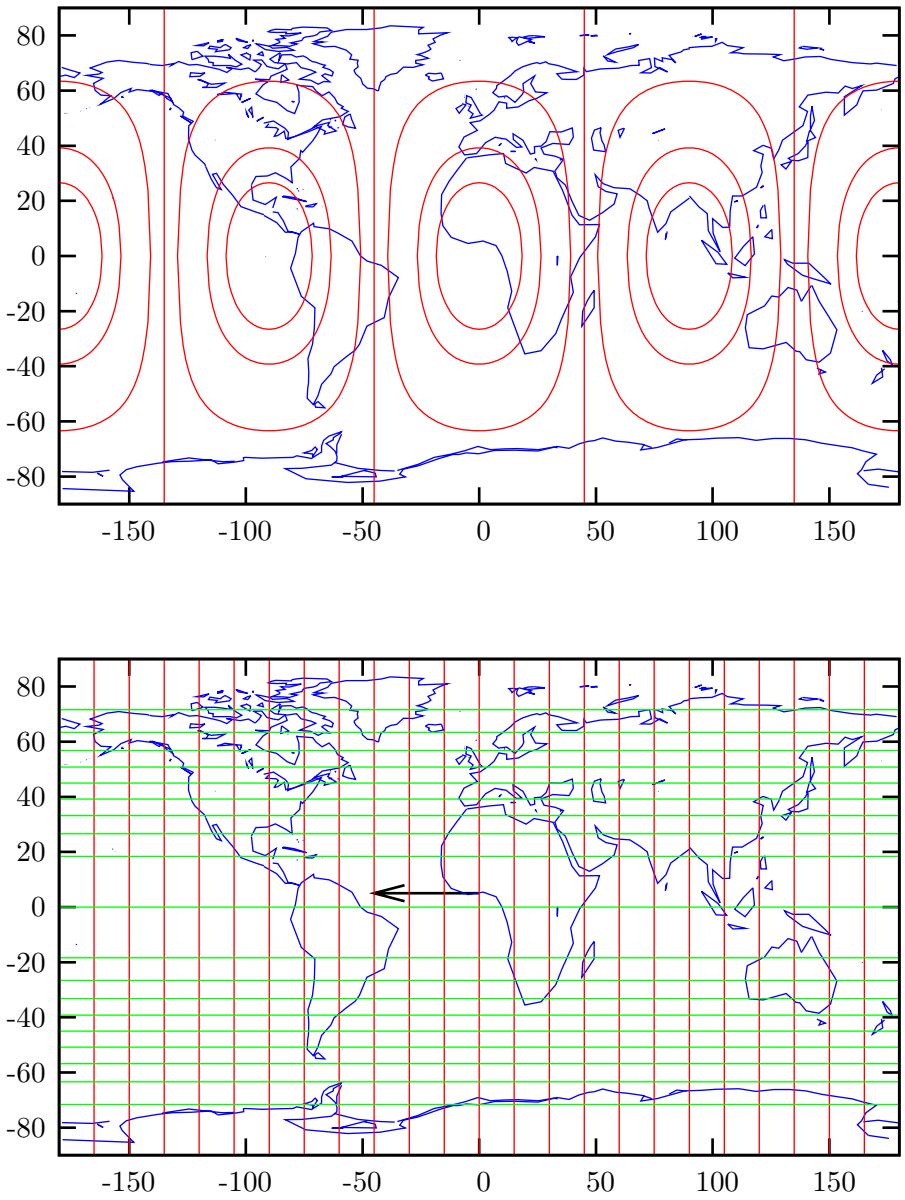


Figure 12: Equipotential lines (top) and cotidal, green, and corange, red, lines (bottom) for the semi-diurnal potential. The arrow indicates the direction of increasing cotidal time.

5.2 Variation of the declination of the tide generating body

The angle δ in figure 9b and (7) varies in time, based on the location of the tide generating body relative to the plane of the equator. This variation has a time period related to the precession of the equinoxes for the Sun and to the precession of the lunar node for the Moon.

5.2.1 Precession of the equinoxes

The Earth's rotational axis is tilted, at present at $24°27'$, to the ecliptic.[4] As the Earth rotates about the Sun, this means that the declination of the Sun varies, as shown in figure 13a. Hence only after a 'year' is the declination expected to return to its initial value.

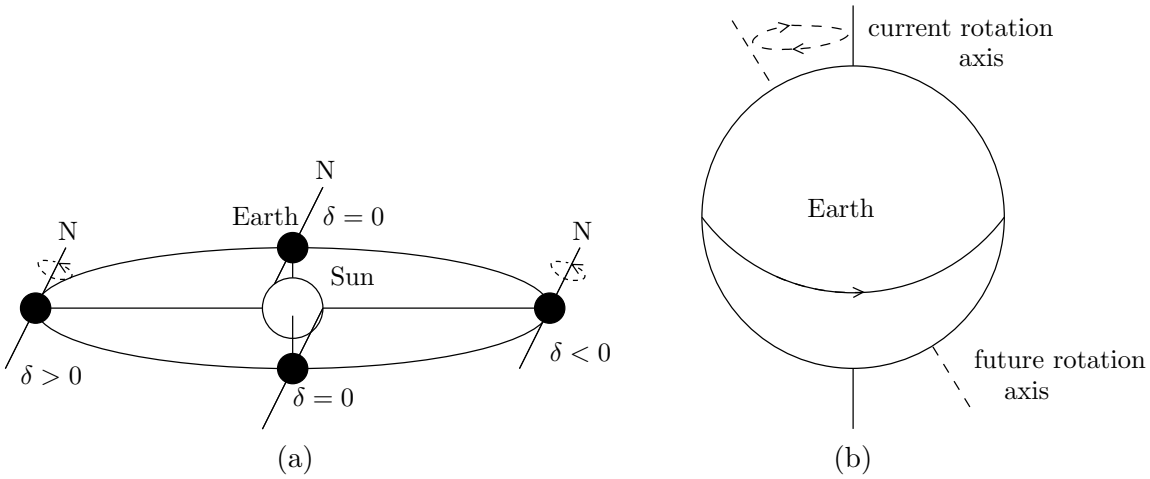


Figure 13: (a) The variation in declination due to the rotation of the Earth about the Sun. (b) The precession of the equinoxes.

Lunar and solar gravity act on the oblate Earth, making it spin like a top, with its rotation axis precessing as depicted in figure 13b. The celestial point where the Sun crosses the plane of the equator, moving from south to north is known as the *Point of Aries* or the vernal equinox and is the point Υ in figure 9 at which $\delta = 0$. Due to precession, it moves eastward relative to the fixed stars.[5] The period of precession is 25570 years. The time period of the revolution of the Earth about the Sun from vernal equinox to vernal equinox is the tropical year and is 365.242199 mean solar days. This is the 'year' we are interested in for the declination returning to its original value.

5.2.2 Precession of the lunar node

The plane of the Moon's orbit around the Earth is inclined at $5°08'$ with respect to the ecliptic. It precesses with a period of 18613 years due to the Earth's gravity.

[4] The plane of the Earth's orbit around the Sun.

[5] In antiquity, it *was* in the constellation Aries. Now it is in Pisces.

The intersection of the plane of the Moon's orbit with the equatorial plane as the Moon goes from south to north is the *ascending lunar node*, Ω. The mean time period separating adjacent passages through Ω is the tropical month of 27.321582 mean solar days. This is the time period before the same declination of the Moon is again achieved.

5.3 Variation of the orbital distance of the TGB

In reality, R is not fixed in (7), since the Moon and Sun are not a constant distance from the Earth. Here we consider how the orbital distance of the Moon varies in time. The same analysis also applied for the Sun.

5.3.1 Kepler's laws

The orbit of the Moon about the Earth is an ellipse (Kepler's first law), as we now show. Consider the geometry shown in figure 14. From Newton's laws,

$$m\ddot{\mathbf{x}} = -GmM\mathbf{x}/R^3,$$

where $\mathbf{x}$ is the position vector of the Moon relative to the Earth.

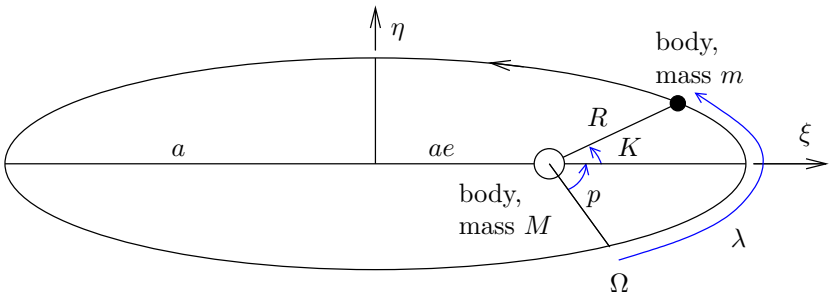


Figure 14: The Moon of mass m in orbit about the Earth of mass M. Assume that the Earth is fixed in space.

Rewriting this equation in polar coordinates, with $x = R\cos\lambda_e$ and $y = R\sin\lambda_e$ gives

$$\ddot{R} - R\dot{\lambda_e}^2 = -GM/R^2, \qquad R\ddot{\lambda_e} + 2\dot{R}\dot{\lambda_e} = 0.$$

Integrating the second equation gives Kepler's second law

$$R^2\dot{\lambda_e} = \text{constant} = h, \tag{8}$$

which says that the line joining the orbiting Moon and the Earth sweeps out equal areas in equal intervals of time. Integrating the first equation, by setting $u = 1/R$, gives

$$\frac{1}{R} = A'\cos K + \frac{GM}{h^2}, \tag{9}$$

where A' is a constant of integration and we have used the fact that $1/R$ is symmetric about the line $\eta = 0$. This is the equation of an ellipse with semi-major axis a and eccentricity e satisfying $A' = e/a(1-e^2)$ and $gM/h^2 = 1/a(1-e^2)$.

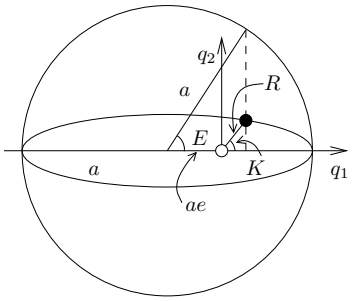


Figure 15: The definition of the eccentric anomaly E.

We now want to see how this modifies the potential given in (7). Instead of using the *true anomaly*, K, it is more useful to write (9) in terms of the *eccentric anomaly*, E, defined as shown in figure 15. Then it is possible to write

$$R = a(1 - e\cos E) \tag{10}$$

and the motion of the Moon in its orbit is given by *Kepler's equation*

$$E - e\sin E = \omega_k t, \tag{11}$$

in which $t = 0$ at perigee[6] and $\omega_k = \sqrt{GM/a^3}$, the *Kepler frequency*. The *mean anomaly*, E_0 is related to E by

$$E_0 = E - \sin E$$

and increases uniformly in time: $E_0 = \omega_k t$ from (11).

In (7) we then have, assuming the eccentricity e is small,

$$R_0/R \equiv a/R = (1 - e\cos E)^{-1} \sim 1 + e\cos E_0 + e^2\cos 2E_0 + \ldots,$$

from (10) and

$$R_0/R = a/R = (1 + e\cos K)/(1 - e^2),$$

from (9). Hence $K \simeq E_0 + 2e\sin E_0 + \ldots$.

Remembering from figure 14 that the *ecliptic longitude*, $\lambda_e = p + K$ and setting the *mean longitude* to be $h = p + E_0$ we obtain

$$\begin{aligned} R/R_0 &= 1 + e\cos(h - p) + e^2\cos 2(h - p) + \ldots, \\ \lambda_e &= h + 2e\sin(h - p) + \ldots. \end{aligned}$$

This can be simply translated into geographical coordinates for the Sun as the tide generating body. For the Sun

$$\sin\delta = \sin\lambda_e \sin\epsilon, \qquad A = \lambda_e - \tan^2(\epsilon/2)\sin\lambda_e,$$

[6]The point in the orbit of the Moon nearest the Earth.

where δ is the declination, λ_e is the ecliptic longitude, A is the right ascension of the Sun and ϵ is the angle between the ecliptic and the equatorial plane.

For the Earth–Moon system it is actually more complex to write the solution in terms of the geographical coordinates as the lunar node does not coincide with the point of Aries and the Moon's orbit is not in the ecliptic. Furthermore, there is a strong solar perturbation to its orbit.

5.4 Tidal harmonics

The effect of the variations of declination and distances to the tide generating bodies is to alter the coefficients for terms in (7). These variations may be Fourier decomposed and result in modulations of the basic tidal frequencies: the long-period, diurnal and semi-diurnal. Then the potential given in (7) may be written as

$$V(\lambda,\theta) = V_0(\lambda,\theta) + V_1(\lambda,\theta) + V_2(\lambda,\theta), \tag{12}$$

where

$$V_s(\lambda,\theta) = DG_s \sum_j C_j \cos(\sigma_j t + s\lambda + \theta_j)$$

with $G_0 = (1 - 3\sin^2\theta)/2$, $G_1 = \sin 2\theta$ and $G_2 = \cos^2\theta$; D the Doodson constant and C_j the amplitude of the component. The harmonic frequency σ_j is a linear combination of the angular velocity of the Earth's rotation ω [already seen in (7) in the hour angle] and the sum and the difference of angular velocities ω_k with $k = 1,\ldots,5$ which are the five fundamental astronomical frequencies, having the largest effect modifying the potentials (it is possible to include many more). These five frequencies are given in table 1. Hence

$$\sigma_j = s\omega + \sum_{k=1}^{5} m_k^j \omega_k,$$

where $s = 0, 1, 2$ for the long-period, diurnal and semi-diurnal respectively; $m_k^j = 0, \pm1, \pm2, \ldots$ and ω is either taken to be $\omega_0 - \omega_1$ for the Moon as the TGB, or $\omega_0 - \omega_2$ for the Sun with ω_0 the sidereal[7] frequency. λ is the longitude of the observer.

All tidal harmonics with amplitudes $C > 0.05$ are given in table 2.

5.4.1 Doodson numbers

For convenience, the frequencies σ_j may be written as

$$\text{Doodson number} = s\, m_1^j\, m_2^j\, m_3^j\, m_4^j\, m_5^j + 0\,5\,5\,5\,5\,5,$$

where the addition of 055555 is simply so that the Doodson numbers are all positive (since in general the m_i^j lie in the range $-5 \le m_i^j < 5$). The Doodson numbers are also given in table 2.

[7]The length of time between consecutive passes of a given 'fixed' star in the sky over the Greenwich meridian. The sidereal day is 23 hr56 min, slightly shorter than the 'normal' or solar day because the Earth's orbital motion about the Sun means the Earth has to rotate slightly more than one turn with respect to the 'fixed' stars in order to reach the same Earth–Sun orientation.

Period	Nomenclature
$360^\circ/\omega_1 = 27.321582\,\text{days}$	period of lunar declination
$360^\circ/\omega_2 = 365.242199\,\text{days}$	period of solar declination
$360^\circ/\omega_3 = 8.847\,\text{years}$	period of lunar perigee rotation
$360^\circ/\omega_4 = 18.613\,\text{years}$	period of lunar node rotation
$360^\circ/\omega_5 = 20940\,\text{years}$	period of perihelion rotation

Table 1: The fundamental periods of the Earth's and the Moon's orbital motion. *From Bartels* [2].

5.4.2 Spectra of the tides

Figure 16 shows a plot of the spectrum of equilibrium tides with frequencies near twice per day (the semi-diurnal tides). The spectrum is split into groups separated by a cycle per month (0.55°hr^{-1}). Each of these is further split into groups separated by a cycle per year (0.04°hr^{-1}). The finest splitting in the figure is at a cycle per 8.847 years ($0.0046^\circ\text{hr}^{-1}$).

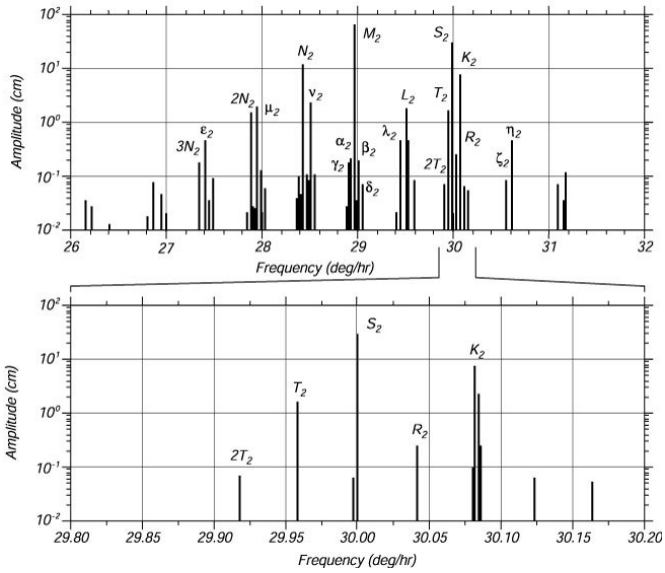


Figure 16: A spectrum for equilibrium tides. *From* `oceanworld.tamu.edu/resources/`.

Notes by Josefina Arraut and Anja Slim.

References

[1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).

[2] J. Bartels, *Handbuch der Physik* (Springer, Berlin, 1967), Vol. 48.

Ampl., C	Frequency, σ $^\circ\,\mathrm{hr}^{-1}$	Period $360^\circ/\sigma$	Doodson number	Notation
		Long period tides		
0.2341	0		055555	S_0 (solar constant)
0.5046	0		055555	M_0 (lunar constant)
0.0655	$\omega_4 = 0.00221$	18.613 years	055565	– (nodal M_0)
0.0729	$2\omega_2 = 0.08214$	182.621 days	057555	S_{sa} (declinational S_0)
0.0825	$\omega_1 - \omega_3 = 0.54437$	27.555 days	065455	M_m (elliptical M_0)
0.1564	$2\omega_1 = 1.09803$	13.661 days	075555	M_f (declinational M_0)
0.0648	$2\omega_1 + \omega_4 = 1.10024$	13.663 days	075565	– (nodal M_0)
		Diurnal tides		
0.0722	$\omega_(- 2\omega_1 + \omega_3 = 13.39866$	26.868	135655	Q_1 (elliptical O_1)
0.0710	$-\omega_(+ \omega_1 - \omega_4 = 13.94083$	25.823	145565	– (nodal O_1)
0.3769	$\omega_(- \omega_1 = 13.94304$	25.819	145555	O_1 (basic lunar)
0.1755	$\omega_\odot - \omega_2 = 14.95893$	24.066	163555	P_1 (basic solar)
0.1682	$\omega_\odot + \omega_2 = 15.04107$	23.934	165555	K_1^S (declinational P_1)
0.3623	$\omega_(+ \omega_1 = 15.04107$	23.934	165555	K_1^M (declinational O_1)
0.0718	$\omega_(- \omega_1 + \omega_4 = 15.04328$	23.931	145565	– (nodal K_1^M)
		Semi-diurnal tides		
0.1739	$2\omega_(- \omega_1 + \omega_3 = 28.43973$	12.658	245655	N_2 (elliptical M_2)
0.9081	$2\omega_(= 28.98410$	12.421	255555	M_2 (basic lunar)
0.4229	$2\omega_\odot = 30.00000$	12.000	273555	S_2 (basic solar)
0.0365	$2\omega_\odot + 2\omega_2 = 30.08214$	11.967	275555	K_2^S (declinational S_2)
0.0786	$2\omega_(+ 1\omega_2 = 30.08214$	11.967	275555	K_2^M (declinational M_2)
		Combined tides		
0.5305	$\omega_0 = 15.04107$	23.934		K_1 (lunar-solar declinational)
0.1151	$2\omega_0 = 30.08214$	11.967		K_2 (lunar-solar declinational)

Table 2: The tidal harmonics with amplitude coefficients $C > 0.05$. $\omega_\odot$ is $\omega_0 - \omega_2$, associated with the Sun and $\omega_($ is $\omega_0 - \omega_1$, associated with the Moon. In the table, for the Doodson numbers, $\omega = \omega_($ and hence $\omega_\odot = \omega_(+ \omega_1 - \omega_2$. Note that the table also includes the K_2^S harmonic, even though it's amplitude is less than 0.05. Its frequency coincides with that of K_2^M and they are completely indistinguishable. They are combined into the single lunar-solar semi-diurnal wave, K_2. *From Bartels* [2].

Lecture 2: The Role of Tidal Dissipation and the Laplace Tidal Equations

Myrl Hendershott

1 Introduction

In this lecture we make a first attempt to describe the energetics of tides. We first provide some discussion of their influence on the global processes of the earth by relating tidally induced dissipation to the change in the earth's rotation rate, and consider whether there is any fossil evidence of such a change. We then approach this question of dissipation a little more formally, from the perspective of an angular momentum budget of the earth-moon system. Finally, we develop the dynamics of tides from first principles, starting with the Navier-Stokes equations on a rotating planet and finally obtaining the Laplace tidal equations.

2 Energetic Dissipation

We shall first consider the relationship between tidal dissipation and the rotation rate of the earth. If we neglect the internal dynamics of the earth, regarding it as a collection of processes that will eventually lead to dissipation, then the persistent energy of the planet is due to rigid-body rotation, whose rate of change is

$$E_t = \frac{\partial}{\partial t}\left(\tfrac{1}{2}C\Omega^2\right), \tag{1}$$

where C is the earth's moment of inertia along the polar axis and Ω is the earth's rotation rate. The change in Ω is then given by

$$\Omega_t = \frac{E_t}{C\Omega}. \tag{2}$$

Historically, astronomical data is used to infer Ω_t, which is then used to compute E_t. But to help motivate this relationship, we would like to compare it to some sort of observational record. For the moment, let us suppose that we have some rough estimate for the tidal dissipation. We may then use this to demonstrate how this lead to a variation in the length of day. Such variation could then leave an imprint in the fossil records, for example. So in

this case, if we use the following values,[1]

$$\begin{aligned} E_t &= -4.0 \times 10^{19} \text{ erg sec}^{-1}, \\ C &= 8.043 \times 10^{44} \text{ g cm}^2, \\ \Omega &= \frac{2\pi}{\text{the sidereal period}} = \frac{2\pi}{86164 \text{ sec}} \\ &= 7.292 \times 10^{-5} \text{ sec}^{-1}, \end{aligned}$$

then we would presume that the earth's rotation rate is currently decreasing at about

$$\Omega_t = -6.8 \times 10^{-22} \text{ rad sec}^{-2}.$$

We can use this to estimate the variation in length of day (LOD),[1]

$$\begin{aligned} \Delta(\text{LOD}) &= \tau' - \tau \\ &= 2\pi \left(\frac{1}{\Omega'} - \frac{1}{\Omega} \right), \end{aligned}$$

where

$$\Omega' = \Omega + \Delta\Omega \simeq \Omega + \Omega_t \Delta t,$$

so that

$$\begin{aligned} \Delta(\text{LOD}) &\simeq -\left(\frac{2\pi}{\Omega} \right) \left(\frac{\Omega_t}{\Omega} \right) \Delta t \\ &= 6.9 \times 10^{-8} \text{ sec} \end{aligned} \tag{3}$$

for $\Delta t = 1$ day $= 86400$ sec. Or, in more appropriate units,

$$\Delta(\text{LOD})/\text{day} = 2.5 \text{ msec cy}^{-1}.$$

As long as $\Omega_t \Delta t \ll \Omega$ remains a reasonably accurate statement, it should be possible to extrapolate about the LOD over epochal times. If we take $-\Delta t$ to be 400 million years, then $\Omega_t \Delta t \approx 0.12\Omega$ and our estimate for a constant LOD variation should be accurate within about 10%. Then in we compare the difference between current years and 400 million years ago, we find the following:

Today	365 days per year
400 Million Years Ago	414 days per year

Although mass factor has been neglected, this does demonstrate that there has likely been a significant change in both the dynamical and radiation cycle of the earth. A natural question is whether there is any evidence supporting the conjecture that the LOD was longer in the past, and how much Ω_t and E_t may have varied. The fossil record offers a possibility, since biological activity should be sensitive to rotation induced variations in the radiation cycle; this is explored in the next section.

[1]The angular rotation per day is slightly greater than 2π, but the error is insignificant for our purposes.

3 Biological Records

Several groups of organisms leave records in the skeletal parts of their accruing tissue, in the form of sequential and repetitive layers. These layers are interpreted in growth increments, and the sequence of layers appears to be a consequence of modulation of growth by internal rhythms inherent to the animal and environmental conditions.

Growth patterns that are controlled by astronomical phenomena are of particular interest. A comparison of frequencies found in living and fossil specimens may reveal indications of the constancy or periodicity of certain astronomical phenomena. One of the types of organisms that have been studied in a geophysical context is corals.

Seasonal fluctuations in the rate of coral growth were first reported by R. P. Whitefield in 1898, who described undulations on some surfaces of living corals and suggested that these represented annual growth increments associated with seasonal water temperature changes. However, the detailed mechanism by which growth occurs and the factors controlling the rate of growth are still inadequately understood.

The skeletal part of corals consist of several elements, one of which, the epitheca, reveals a fine structure of ridges that are parallel to the growing edge (figure 1). These ridges are interpreted as growth increments and they suggest a periodic fluctuation in the rate of calcium carbonate secretion. The rate of deposition of these growth increments in modern reef-forming corals is believed to be daily.

An indication that the growth ridges are daily is that modern corals typically add about 360 such increments per year, suggesting that the solar day controls the frequency of deposition[2, 3]. Although, factors other than variable daylight may be important here in modulating the growth rates, indirect evidence suggests that the solar day has remained the dominant periodicity in corals studied; Devonian corals studied by Wells[2] show about 400 daily growth increments between successive seasonal annulations, in keeping with expected value, if present tidal acceleration of the Earth has remained roughly constant over last $3\text{–}4 \times 10^8$ years. The same qualitative results can be found in molluscs and stromatolites too.

4 Angular Momentum

The consequences of tidal dissipation can also be seen in the receding of the moon from the earth. If we ignore the rotation of the moon and regard it as a point in space moving in a simple circular orbit, and again focus on the rigid-body rotation of the earth, then the total angular momentum of the earth-moon system is conserved and

$$\frac{\partial}{\partial t}\left(C\Omega + ml^2\omega\right) = 0, \tag{4}$$

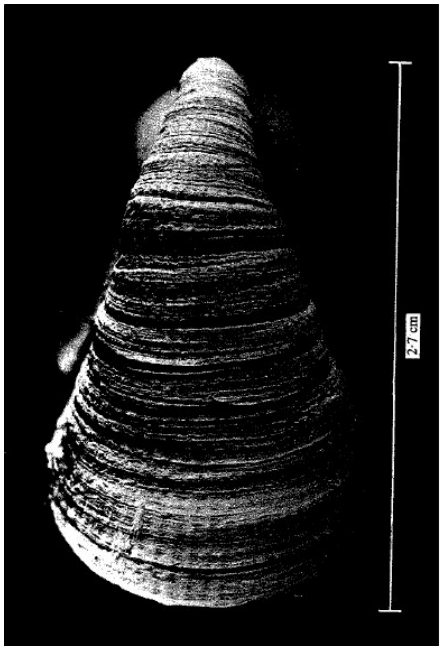


Figure 1: Middle Devonian coral epitheca from Michigan, U.S.A.

where m is the moon's mass, l is the orbital radius, and ω is the orbital frequency. These parameters currently have values of

$$\begin{aligned} m &= 7.35 \times 10^{25} \text{ g}, \\ l &= 3.84 \times 10^{10} \text{ cm}, \\ \omega &= \frac{2\pi}{27.32166 \times 86400 \text{ sec}} = 2.66 \times 10^{-6} \text{ rad sec}^{-1}. \end{aligned}$$

Expansion of the angular momentum equation gives

$$C\Omega_t + m\left(2l\omega l_t + l^2\omega_t\right) = 0.$$

To relate the change in distance to the change in frequency, we note that since the motion is assumed circular,

$$\frac{GMm}{l^2} = m\omega^2 l,$$

so that

$$\omega^2 l^3 = GM \quad \text{(Kepler's Third Law)},$$

and hence

$$\omega_t = -\frac{3}{2}\left(\frac{\omega}{l}\right) l_t. \tag{5}$$

We then obtain an expression for the rate at which the moon drifts from the earth,

$$l_t = -\frac{2C\Omega_t}{ml\omega}. \tag{6}$$

Laser ranging of the moon tells us that

$$l_t = 3.8 \text{ cm yr}^{-1} = 1.2 \times 10^{-7} \text{ cm sec}^{-1}$$

from which we infer that the change in the lunar cycle is

$$\omega_t = -1.25 \times 10^{-23} \text{ rad sec}^{-2} = -25.7'' \text{ cy}^{-2}.$$

This simple estimate is in close agreement with a more sophisticated calculation by Brosche and Sündermann, who obtained a result of $-26.06''$ cy^{-2}.[4]

From (6), we find that the rate of change of the earth's rotation is

$$\Omega_t = -5.6 \times 10^{-22} \text{ rad sec}^{-1}.$$

Based on this rotation rate, the loss of energy of the moon, from (2), is

$$E_t = -3.3 \times 10^{19} \text{ erg sec}^{-1} = -3.3 \text{ TW}$$

which, despite our idealizations, is in fairly good agreement with more sophisticated astronomical calculations of 3.75 ± 0.08 TW for lunar dissipation.[5]

5 Terrestrial Coordinates and the Traditional Approximation

In this section we consider the dynamics of a shallow fluid on a rotating planet. Later sections will introduce further approximations, which will lead us to the Laplace tidal equations (LTE) for each vertical mode.

The Navier-Stokes equations for an incompressible rotating fluid in spherical (terrestrial) coordinates are

$$\frac{D_s u}{Dt} - \left[2\Omega + \frac{u}{r\cos\theta}\right] v \sin\theta + \left[2\Omega + \frac{u}{r\cos\theta}\right] w\cos\theta = -\frac{p_\phi}{\rho r\cos\theta} - \frac{\Phi_\phi}{r\cos\theta} + X, \tag{7a}$$

$$\frac{D_s v}{Dt} + \left[2\Omega + \frac{u}{r\cos\theta}\right] u\sin\theta + \left[\frac{v}{r}\right] w = -\frac{p_\theta}{\rho r} - \frac{\Phi_\theta}{r} + Y, \tag{7b}$$

$$\frac{D_s w}{Dt} - \left[\frac{v}{r}\right] v - \left[2\Omega + \frac{u}{r\cos\theta}\right] u\cos\theta = -\frac{p_r}{\rho} - \Phi_r + Z, \tag{7c}$$

$$\frac{1}{r\cos\theta}\left[u_\phi + (v\cos\theta)_\theta\right] + \frac{1}{r^2}\left(r^2 w\right)_r = 0, \tag{7d}$$

where θ and ϕ are the latitude and longitude, Φ is the geopotential, and $\mathbf{X} = (X, Y, Z)$ represent dissipative forces. The advective derivative, $\frac{D_s}{Dt}$, is

$$\frac{D_s}{Dt} = \frac{\partial}{\partial t} + \frac{u}{r\cos\theta}\frac{\partial}{\partial\phi} + \frac{v}{r}\frac{\partial}{\partial\theta} + w\frac{\partial}{\partial r}$$

and the self-gravitating potential of the earth, including centrifugal forcing, is

$$\Phi = -\frac{GM_E}{r} + A(r,\theta) - \frac{1}{2}\Omega^2 r^2\cos^2\theta, \tag{8}$$

where $A(r,\theta)$ describes the deviation of gravitation from that of a spherical earth.

The centrifugal forcing of the earth causes the earth to deform, so that the equilibrium shape is closer to an oblate spheroid than a perfect sphere. But since the geopotentials are also deformed to match the shape of the earth, this tends to produce a situation that is only a slight distortion of the purely spherical dynamics. Veronis [6] has shown that the actual dynamics can be explicitly written as a perturbation series in terms of the ellipticity of the earth, e, with the leading order equations corresponding to dynamics on a sphere of radius $a = \frac{1}{2}(r_{\text{eq}} + r_{\text{pole}})$ and a uniform gravitational acceleration. The errors are more or less bounded by $\frac{3}{2}e \approx \frac{1}{200}$, or 0.5%, so it is generally reasonable to treat the earth as spherical if we restrict ourselves to large-scale flows.

The full set of equations satisfy the usual conservation laws. For example, if we neglect gravitational and dissipative forcing, the kinetic energy of a fluid parcel is balanced by the work due to pressure,

$$\frac{D_s}{Dt}\left[\frac{u^2+v^2+w^2}{2}\right] = -\frac{\mathbf{u}\cdot\nabla p}{\rho} \tag{9}$$

and the angular momentum is balanced by the pressure torque,

$$\frac{D_s}{Dt}\left[r\cos\theta\left(u+\Omega r\cos\theta\right)\right] = -\frac{p_\phi}{\rho}. \tag{10}$$

When integrated over the earth, both quantities are conserved.

From incompressibility $\left(\frac{W}{H} \ll \frac{U}{L}\right)$ and a shallow water aspect ratio ($H \ll L$), we expect a scaling where $W \ll U$ and that it is reasonable to neglect the centrifugal and Coriolis forces involving w and to replace r by a. But these changes also disrupt the conservation of energy and angular momentum unless we also neglect the centrifugal and Coriolis forces in z. This so-called *traditional approximation* is not always justifiable from scale analysis for flows of a homogeneous fluid shell , but it does produce a set of equations that is generally consistent with terrestrial fluid flow.

After applying the traditional approximation, we have

$$\frac{D_{sa}u}{Dt} - \left[2\Omega + \frac{u}{a\cos\theta}\right] v\sin\theta = -\frac{p_\phi}{\rho a\cos\theta} + X, \tag{11a}$$

$$\frac{D_{sa}v}{Dt} + \left[2\Omega + \frac{u}{a\cos\theta}\right] u\sin\theta = -\frac{p_\theta}{\rho a} + Y, \tag{11b}$$

$$\frac{D_{sa}w}{Dt} = -\frac{p_z}{\rho} - g + Z, \tag{11c}$$

$$\frac{1}{a\cos\theta}\left(u_\phi + (v\cos\theta)_\theta\right) + w_z = 0, \tag{11d}$$

where $z = r - a$ is the displacement from the earth's surface, D_{sa}/Dt is the advective derivative with $r = a$, and $g = GM_E/a^2$ is the radial acceleration due to the (now spherical) geopotentials.

6 Boussinesq Approximation

Because density variations are expected to be small, it is appropriate to introduce a Boussinesq approximation, where we only consider small variations from a hydrostatic basic state. Let p_0 denote the hydrostatic pressure, and also let $\overline{\rho}_0$ be the mean density and $\rho_0(z)$ the steady state variation from $\overline{\rho}_0$. Then since $\mathbf{u} = 0$ and $\mathbf{X} = 0$ at equilibrium, $p_0 = p_0(z)$ and

$$\frac{d\overline{p}}{dz} = -\left(\overline{\rho}_0 + \rho_0\right) g. \tag{12}$$

So if $p = p_0 + p'$ and $\rho = \overline{\rho}_0 + \rho_0 + \rho'$, then the hydrostatic balance is removed and, to leading order in density, the equations are

$$\frac{Du}{Dt} - \left[2\Omega + \frac{u}{a\cos\theta}\right] v \sin\theta = -\frac{p'_\phi}{\overline{\rho}_0 a\cos\theta} + X, \tag{13a}$$

$$\frac{Dv}{Dt} + \left[2\Omega + \frac{u}{a\cos\theta}\right] u \sin\theta = -\frac{p'_\theta}{\overline{\rho}_0 a} + Y, \tag{13b}$$

$$\frac{Dw}{Dt} = -\frac{p'_z}{\overline{\rho}_0} - g\frac{\rho'}{\overline{\rho}_0} + Z, \tag{13c}$$

$$\frac{1}{a\cos\theta}\left(u_\phi + (v\cos\theta)_\theta\right) + w_z = 0, \tag{13d}$$

$$\frac{D}{Dt}\left(\rho_0 + \rho'\right) = 0, \tag{13e}$$

where the thermodynamic incompressibility equation has been written explicitly, and the advective derivative subscript has been dropped.

Since the study of tidal dynamics focuses on the generation and propagation of gravitationally forced tidal waves, we will use the linearized Boussinesq equations. If we also assume that the perturbation flow is hydrostatic, then the system of equations for freely propagating waves is

$$u_t - (2\Omega\sin\theta)\, v = -\frac{p_\phi}{\overline{\rho}_0 a\cos\theta}, \tag{14a}$$

$$v_t + (2\Omega\sin\theta)\, u = -\frac{p_\theta}{\overline{\rho}_0 a}, \tag{14b}$$

$$0 = -\frac{p_z}{\overline{\rho}_0} - g\frac{\rho}{\overline{\rho}_0}, \tag{14c}$$

$$\frac{1}{a\cos\theta}\left(u_\phi + (v\cos\theta)_\theta\right) + w_z = 0, \tag{14d}$$

$$\rho_t = \frac{\overline{\rho}_0}{g} N^2 w, \tag{14e}$$

where the buoyancy frequency is

$$N(z) = \left(-\frac{g}{\overline{\rho}_0}\frac{d\rho_0}{dz}\right)^{\frac{1}{2}}$$

and primes have been dropped. The vertical velocity w becomes a diagnostic variable that can be computed from the density (thermodynamic) equation.

The original system was assumed to be incompressible. If we had included compressible effects, then the only major difference under the Boussinesq approximation would be that the buoyancy frequency is

$$N(z) = \left(-\frac{g}{\overline{\rho}_0} \frac{d\rho_0}{dz} - \frac{g^2}{c^2} \right)^{\frac{1}{2}},$$

where c is the speed of sound.

7 Vertical Mode Decomposition

Under certain situations, such as in the absence of a mean flow and with flat topography, we can decompose our solutions into a set of *vertical modes*, where each mode corresponds to the flow of a shallow water fluid. This provides an interpretation that relates the stratified continuum to a sequence of layered fluids, and isolates the barotropic waves (the unstratified dynamics) from the baroclinic waves. The analysis here closely follows Pedlosky.[7]

For a flat bottom (at, say, $z = -D_*$) we can separate the variables of the problem into a function of z and a function of horizontal and time variables so that

$$\begin{aligned} u &= U(\phi, \theta, t)F(z), \\ v &= V(\phi, \theta, t)F(z), \\ w &= W(\phi, \theta, t)G(z), \\ p &= \overline{\rho}_0 g \zeta(\phi, \theta, t)F(z). \end{aligned}$$

When we insert these expressions into the equations of motion, the horizontal momentum equations become

$$\begin{aligned} U_t - fV &= -\frac{g\zeta_\phi}{a\cos\theta}, \\ V_t + fU &= -\frac{g\zeta_\theta}{a}. \end{aligned}$$

Now let us apply these forms to the continuity equation,

$$\frac{U_\phi}{a\cos\theta} + \frac{(V\cos\theta)_\theta}{a\cos\theta} = -W\frac{G_z}{F}.$$

All terms except the ratio $\left(\frac{G_z}{F}\right)$ are independent of z while each term of this ratio is a function only of z. The only way this can hold for every z is if both sides equal a constant. Let us define this constant as

$$\frac{G_z}{F} \equiv \frac{1}{h}.$$

Then the continuity equation becomes

$$\frac{U_\phi}{a\cos\theta} + \frac{(V\cos\theta)_\theta}{a\cos\theta} + \frac{W}{h} = 0.$$

Applying the separable forms to the adiabatic equation yields

$$\zeta_t + W\frac{G}{F_z}\frac{N^2}{g} = 0,$$

which becomes

$$\zeta_t + W\frac{G}{G_{zz}}\frac{N^2}{gh} = 0.$$

Again we see that because ζ and W are not functions of z the coefficient of W must be constant. We can choose this constant to be -1 without any loss of generality (a different constant will only change the definition of h). With this choice the adiabatic equation becomes

$$\zeta_t = W.$$

This choice yields an equation for G,

$$G_{zz} + \frac{N^2}{gh}G = 0.$$

This is a homogeneous differential equation with, generally, non-constant coefficients since N is a function of z and there is a free parameter h. The problem is not complete until the boundary conditions are established. In order to have w vanish on $z = -D_*$ we must take

$$G(z) = 0, \quad z = -D_*.$$

At the free surface with a shallow water assumption the conditions are that the free surface displacement, which here we will call z_T satisfies

$$w = WG(z_T) = \frac{\partial z_T}{\partial t}.$$

While the total pressure is atmospheric pressure, which we will take to be a constant (zero), thus

$$P_{\text{total}} = p_0(z_T) + g\zeta F(z_T) = p_0(0) + \frac{dp_0}{dz}z_T + \ldots + \rho_0 g\zeta F(0).$$

Now keeping only linear terms let us derive the former equation with respect to time and combine it with the linearized kinematic equation,

$$\begin{aligned} 0 &= \frac{dp_0}{dz}\frac{\partial z_T}{\partial t} + \rho_0 g\zeta_t F(0), \\ 0 &= -\rho_0 gWG(0) + \rho_0 g\zeta_t F(0), \\ \rho_0 gWG(0) &= \rho_0 g\zeta_t F(0). \end{aligned}$$

But, from the continuity equation we know that,

$$\zeta_t = W,$$

thus,

$$G(0) = F(0) = hG_z(0).$$

Which means that the final condition for G is,

$$G_z - G/h = 0, \quad z = 0.$$

Let us summarize the equations for the resulting eigenvalue problem,

$$\begin{aligned} G_{zz} + \tfrac{N^2}{gh} G &= 0, \\ G &= 0, \qquad z = -D_*, \\ G_z - G/h &= 0, \qquad z = 0. \end{aligned}$$

Using the relations between F and G we obtain as an equally valid alternative problem,

$$\begin{aligned} \left(\tfrac{F_z}{N^2}\right)_z + \tfrac{1}{gh} F &= 0, \\ F_z &= 0, \qquad z = -D_*, \\ F_z - \tfrac{N^2}{g} F &= 0, \qquad z = 0. \end{aligned}$$

The advantage of the second formulation is that the eigenvalue h is not in the boundary condition. These equations can either be solved numerical for a varying N or examined analytically for the case of a constant N.

7.1 Vertical Modes For Constant N

Let us derive h for the case of a constant N. In this case the solution for $G(z)$, which satisfies the boundary condition at $z = -D_*$ is

$$G = A \sin m(z + D_*), \quad m^2 \equiv \tfrac{N^2}{gh},$$

where m is the vertical wavenumber of the solution. The eigenvalue relation for h is obtained from the boundary condition at $z = 0$, and yields,

$$m \cos(md) - \frac{1}{h} \sin(md) = 0,$$

or,

$$\tan(md) = mh = \frac{N^2}{gm},$$

or,

$$\tan(md) = \frac{N^2 d}{g} \frac{1}{md}.$$

We note that

$$\frac{N^2 d}{g} = -\frac{d\frac{d\rho_0}{dz}}{\rho_0} \simeq \frac{\Delta\rho_0}{\rho_0} \ll 1.$$

Thus, the roots of the dispersion relation split into two classes. The first class has roots for which md is order $O(1)$. In that case the right hand side of the dispersion relation is essentially zero and the solutions correspond to the zeros of the tangent function,

$$md = j\pi, \quad j = 1, 2, 3, \ldots.$$

In this approximation, there is an infinite number of roots corresponding to

$$m = -\frac{j\pi}{d}.$$

Therefore, from the definition of m we get,

$$h_j = -\frac{N^2 d^2}{g j^2 \pi^2}.$$

From this we can easily see that the modal structures are,

$$\begin{aligned} G_j &= \sin \tfrac{j\pi z}{d}, \quad j = 1, 2, 3, \ldots, \\ F_j &= \cos \tfrac{j\pi z}{d}, \quad j = 1, 2, 3, \ldots. \end{aligned}$$

Now let us consider the second class for which md is not of order $O(1)$ but, rather, $md \to 0$. Then, the dispersion relation becomes,

$$\tan(md) \simeq md = \frac{N^2 d}{g} \frac{1}{md}.$$

Therefore, from the definition of m we get,

$$h_0 = d,$$

which gives us the barotropic mode of zero vertical velocity with $m_0 \ll 1$ due to the almost no variation of F_0 in depth.

7.2 Vertical Modes For Slowly Varying N

Instead of assuming that stratification remains constant, a more realistic assumption would be to assume that N is slowly varying in some sense, for example if $dN/dz \ll N/l$ for some length scale l, such as the characteristic mode depth. If this remains true, then we can find an asymptotic solution using WKB methods.

If we rescale our equation in terms of a slow variable $Z = \epsilon z$, then the field equation for G is

$$\epsilon^2 G_{ZZ} + \frac{[N(Z/\epsilon)]^2}{g D_n} G = 0. \tag{15}$$

Since N is approximately constant, we have a WKB solution of the form

$$G \sim \exp\left[\frac{1}{\epsilon} S_0(Z) + S_1(Z)\right].$$

After substitution into the differential equation and matching powers of ϵ, we find that

$$S_0(Z) = \pm \frac{i}{\sqrt{g D_n}} \int_{-\epsilon D_*}^{Z} N(Z'/\epsilon) dZ', \tag{16a}$$

$$S_1(Z) = -\frac{1}{2} \ln\left[N(Z/\epsilon)\right], \tag{16b}$$

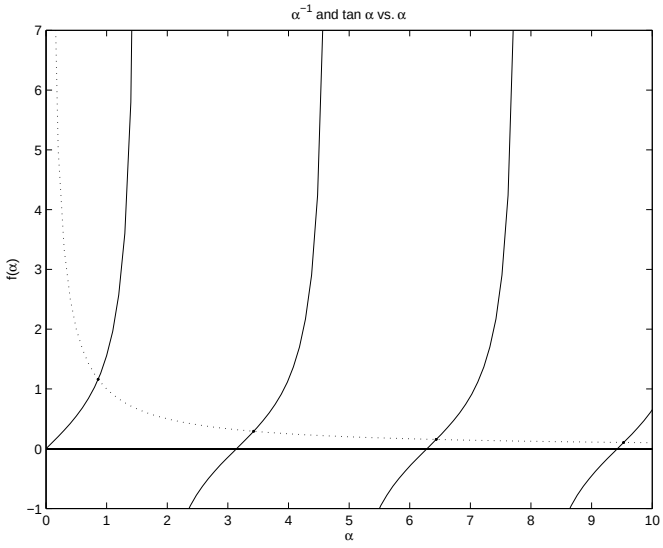


Figure 2: The solid line is the curve for $\tan\alpha$ and the dotted line is for α^{-1}. The intersection of the curves correspond to the roots of the equation $\alpha^{-1} = \tan\alpha$. As α increases, we see that the roots approach $\alpha = \pi n$.

so that G, in terms of z, becomes

$$G(z) \simeq \frac{1}{\sqrt{N(z)}} \exp\left[\pm \frac{i}{\sqrt{gD_n}} \int_{-D_*}^{z} N(z')dz'\right],$$

where by $\pm$, we mean that G is some linear combination of each solution.

We must now apply the boundary conditions to obtain a complete solution. From the surface condition, $G(z) = 0$ at $z = -D_*$, we see that

$$G(z) \simeq \frac{1}{\sqrt{N(z)}} \sin\left[\frac{1}{\sqrt{gD_n}} \int_{-D_*}^{z} N(z')dz'\right]. \tag{17}$$

From the upper condition, $G_z - G/D_n = 0$ at $z = 0$, and using the fact that dN/dz is small, we find that

$$N\sqrt{\frac{D_n}{g}} = \tan\left[\frac{1}{\sqrt{gD_n}} \int_{-D_*}^{0} N(z')dz'\right]. \tag{18}$$

If we let $\alpha_n = \sqrt{\frac{g}{N^2 D_n}}$, then solving for D_n is equivalent to solving for α_n in the equation

$$\frac{1}{\alpha_n} = \tan\left(k\alpha_n\right), \tag{19}$$

where $k = \frac{N}{g}\int_{-D_*}^{0} N(z')dz'$. The solutions to this equation for $k = 1$ are illustrated in figure 2. We can see that the roots correspond to $\alpha_n = \pi n/k$ as n becomes large so that

$$D_n \simeq \frac{\left[\int_{-D_*}^{0} N(z')dz'\right]^2}{g\pi^2 n^2}. \tag{20}$$

Although this is only formally correct for large n, the error is often small even for $n = 1$, and the accuracy only improves with increasing n, so that the WKB solution can usually be safely applied to the entire baroclinic spectrum.

8 Laplace Tidal Equations

From the above equations and the modal solution with respect to the z axis we can derive the Laplace tidal equations,

$$U_t - fV = -\frac{g\zeta_\phi}{a\cos\theta}, \tag{21a}$$

$$V_t + fU = -\frac{g\zeta_\theta}{a}, \tag{21b}$$

$$\frac{U_\phi}{a\cos\theta} + \frac{(V\cos\theta)_\theta}{a\cos\theta} + \frac{W}{h_j} = 0, \tag{21c}$$

$$\zeta_t = W. \tag{21d}$$

And if the two last equations are combined,

$$U_t - fV = -\frac{g\zeta_\phi}{a\cos\theta}, \tag{22a}$$

$$V_t + fU = -\frac{g\zeta_\theta}{a}, \tag{22b}$$

$$\zeta_t + h_j\left(\frac{U_\phi}{a\cos\theta} + \frac{(V\cos\theta)_\theta}{a\cos\theta}\right) = 0. \tag{22c}$$

This set of equations for each mode was considered for a flat bottom and with no external forcing. The friction and the tidal gravitational potential (TGP) can be introduced by simply decomposing these functions into their vertical modes. But since the TGP does not vary significantly across the shallow depth of the fluid, only the barotropic mode is excited noticeably when the bottom is flat, despite the fact that actual observations show strong baroclinic tides.

Although we will not discuss how to approach the full baroclinic case, we can consider the barotropic mode over a variable topography with a weakly elastic bottom (due to the elasticity of the earth). In this case, let $D(\theta, \phi)$ describe the mean depth of a homogeneous fluid, and δ denote the displacement of the elastic bottom. Then the equations for the barotropic mode are generalized to

$$u_t - (2\Omega\sin\theta)\, v = -\frac{g\,(\zeta - \Gamma/g)_\phi}{a\cos\theta} + \frac{F^\phi}{\rho D}, \tag{23a}$$

$$v_t + (2\Omega\sin\theta)\, u = -\frac{g\,(\zeta - \Gamma/g)_\theta}{a} + \frac{F^\theta}{\rho D}, \tag{23b}$$

$$(\zeta_t - \delta_t) + \frac{1}{a\cos\theta}\left[(uD)_\phi + (vD\cos\theta)\right] = 0, \tag{23c}$$

where Γ is the total tide generating potential. We will refer to these as the LTE. Note that an equation for the dynamics of the elastic bottom must be included to fully describe the system.

Notes by Yaron Toledo and Marshall Ward.

References

[1] K. Lambeck, *The Earth's Variable Rotation* (Cambridge University Press, Cambridge, 1980).

[2] J. W. Wells, "Coral growth and geochronometry," Nature **197**, 448 (1963).

[3] D. J. Barnes, "The structure and formation of growth-ridges in scleractinian coral skeletons," Proc. Roy. Soc. London **13**, 1 (1973).

[4] P. Brosche and J. Sündermann, in *Scientific Applications of Lunar Laser Ranging*, edited by J. D. Mullholland (D. Reidel, Dordrecht, 1977), pp. 133–141.

[5] L. H. Kantha, "Tides—a modern perspective," Marine Geodesy **21**, 275 (1998).

[6] G. Veronis, "Large scale ocean circulation," Advances in Applied Mathematics **13**, 1 (1973).

[7] J. Pedlosky, *Waves in the Ocean and Atmosphere: Introduction to Wave Dynamics* (Springer-Verlag, New York, 2003).

Lecture 3: Solutions to Laplace's Tidal Equations

Myrl Hendershott

1 Introduction

In this lecture we discuss assumptions involved in obtaining Laplace's Tidal Equations (LTE) from Euler's equations. We first derive an expression for the solid Earth tides. Solutions of LTE for various boundary conditions are discussed, and an energy equation for tides is presented.

Solutions to the Laplace Tidal equations for a stratified ocean are discussed in §2. We obtain expression for solid earth tide in §3. Different models of dissipation are examined in §4. Boundary conditions for LTE's are discussed in §5. In §6 the energy equation for LTE's is derived.

2 The Laplace Tidal equations for a Stratified ocean

To obtain the LTE for stratified ocean we assume pressure is hydrostatic and seek seperable solutions of the form

$$\begin{aligned} u(x,y,z,t) &= U(x,y,t)F_u(z), \\ w(x,y,z,t) &= W(x,y,t)F_w(z), \\ p(x,y,z,t) &= Z(x,y,t)F_p(z). \end{aligned}$$

The LTE for stratified ocean are

$$U_t - fV = -gZ_x \quad , \tag{1}$$

$$V_t - fU = -gZ_y \quad , \tag{2}$$

$$Z_t + D_n(U_x + V_y) = 0 \quad , \tag{3}$$

$$F_{w_{zz}} + \frac{N^2}{gD_n}F_w = 0 \quad , \tag{4}$$

where n is an index for the normal modes in the ocean. These equations are the constant depth LTE but where D_n is the equivalent depth of each mode, and $D_n \neq D_*$, rather

$$D_n = \frac{[\int_{-D_*}^{0} N(z')dz']^2}{gn^2\pi^2}$$

and $N(z)$ is the buoyancy frequency. For the zeroth mode or $n = 0, D_* \approx D$. The eigenvalue problem in (4) can be solved to determine the D_n using the following boundary conditions

$$\begin{aligned} F_w &= 0 \quad \text{at} \quad Z = -D_*, \\ F_w - D_n F_{w_z} &= 0 \quad \text{at} \quad Z = 0. \end{aligned}$$

2.1 Barotropic Solution (n = 0)

The normal mode equations described above and indexed by n an integer can be solved for specific modes. The zeroth order mode is given by $n = 0$ and is also called the barotropic or external mode. It is characterized by a solution which is depth independent. Below we will solve for the barotropic mode without rotation and no variations in the y direction ($f = 0$ and $\partial/\partial y = 0$). If we assume a plane wave solution, solving (1) - (4) gives,

$$\begin{aligned} Z_0 &= a e^{-i\sigma t + ikx} \ , \\ U_0 &= a \left(\frac{gk}{\sigma} \right) e^{-i\sigma t + ikx} \ , \\ \sigma &= \pm k \sqrt{g D_*}. \end{aligned}$$

Figure 1 shows the barotropic solution for velocity. For the semidiurnal tidal frequency, the phase velocity of the zeroth mode wave, given by $c_0 = \sqrt{gD_0} = \sigma/\kappa = 200$ m/s. Since this speed is also given by λ/T, where λ is the horizontal wavelength of the wave (distance from wave crest to wave crest), then $\lambda = 8640$ km. This wave is very fast and very long. For the case of no rotation this wave is dispersionless, but not when $f \neq 0$.

2.2 Baroclinic Solution, Mode 1 (n = 1)

The first baroclinic mode, indexed by $n = 1$ is also called the first internal mode. The rest of the modes for $n > 1$ are also internal modes and have more variation in depth. We can solve (1) - (4) for $n = 1$ without rotation and with no variations in y ($f = 0$ and $\partial/\partial y = 0$) giving the first internal mode,

$$\begin{aligned} Z_n &= a e^{-i\sigma t + ikx} \ , \\ F_{U_1} &= cos(\frac{\pi(z + D_*)}{D_*}) \ , \\ F_{W_1} &= sin(\frac{\pi(z + D_*)}{D_*}) \ , \\ \sigma &= \pm k \sqrt{g D_1}. \end{aligned}$$

The mode one solution is shown in figure 1. For the baroclinic modes, the phase velocity and horizontal wavelength are given by

$$\begin{aligned} c_n = \frac{N_0 D_*}{n\pi} = \frac{(1.45 \times 10^{-3})(4000m)}{n\pi} = \frac{1.85 \text{ m/s}}{n}, \\ \lambda_{tidal} = \frac{80 \text{ km}}{n}. \end{aligned}$$

So the lower modes travel more quickly, and have larger horizontal scales. The M_2 internal tide is found primarily as a mode 1 tide throughout the world's oceans, while higher order modes tend to dissipate nearer their source. Due to recent (last decade) improvements, the M_2 internal tide can now be seen by satellite altimetry. The ocean surface displacement due to internal tides is given by

$$\frac{w_{\text{free surface}}}{w_{\text{interior maximum}}} \approx \frac{N_0^2 D_*}{gn\pi} \approx 3 \times 10^{-4},$$

which means that for an internal tide displacement of isopycnals on the order of 100m (which is quite large but not impossible), the surface expression would be about 30cm, easily resolved by satellite altimetry measurements which have accuracy on the order of a few centimeters.

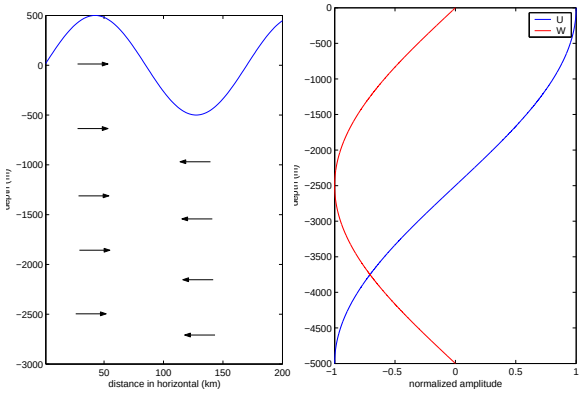


Figure 1: (a) Barotropic or depth independent solution for u velocity. Wave amplitude is greatly exaggerated. (b) Mode 1 solution for u and w velocity.

2.3 Numerical Solution to LTE

The complete problem that we would like to solve numerically to estimate the tides are the stratified linear equations,

$$u_t - fv = -\frac{p_x}{\rho_0} + \Gamma_x, \tag{5}$$

$$v_t + fu = -\frac{p_y}{\rho_0} + \Gamma_y, \tag{6}$$

$$\zeta_t + (uD)_x + (vD)_y = 0, \tag{7}$$

where Γ is full tide generating potential and $D(x, y)$ is the bottom topography. The domain is defined by the coasts of continents, ocean bottom and free surface. However, in order to solve this system of equations one must resolve short horizontal scales due to bottom topography where the bottom boundary condition on w is $w = -\mathbf{u} \cdot \nabla D$. Very few current modes are capable of this, though some have begun to resolve mode 1 in their simulations.

Instead the Laplace tidal equations for u and v may be substituted, of (5)-(7).

$$u_t - fv = -g\zeta_x + \Gamma_x, \tag{8}$$
$$v_t + fu = -g\zeta_y + \Gamma_y, \tag{9}$$

where ζ is the free surface and a smoother bottom topography is substituted,

$$w = -\mathbf{u} \cdot \nabla D_{smooth}. \tag{10}$$

However, because ζ is no longer a function of z while p in (5)-(7) was, we cannot determine $u(z)$ and (8)-(9) will only give the barotropic solution.

In the literature, the TGP is usually neglected and instead the barotropic tides and stratification are specified, which allows the simplification of (5)-(7) as

$$u_t - fv = -\frac{p_x}{\rho_0}, \tag{11}$$
$$v_t + fu = -\frac{p_y}{\rho_0}. \tag{12}$$

From this, the internal tides result from a single scattering of the barotropic tide by bottom relief $w_{int} = u_B \cdot \nabla D(x, y)$. In particular, if we decomposed D into low- and high-passed components,

$$D(x, y) = D_{lo}(x, y) + D_{hi}(x, y). \tag{13}$$

Then the ζ equation and bottom boundary conditions become

$$\zeta_{0t} + \nabla \cdot (\mathbf{u}_B D_{lo}(x, y)) = 0,$$
$$w_{int} = \mathbf{u}_B \cdot \nabla D_{hi}(x, y).$$

And the internal tide results from the bottom topography. However, this neglects multiple scattering from the topography and does not apply when the bottom slope is greater than the characteristic slope of internal waves. Currently, numerical models like the Princeton Ocean Model (POM) solve the (5)-(7). An example of numerically solved tides is shown in figure 2. In this paper, all tidal constituents were solved for using a hydrodynamic model and data assimilation from tide gauges and altimetry [1].

3 Solid Earth Tide

It has long been known that Earth's crust yields elastically to the tidal forces of the moon and sun. If we consider earth to be an incompressible elastic solid, then we can write equations for the deformation of the Earth as the following

$$-p_x + \mu\nabla^2 u = 0, \tag{14}$$
$$-p_z + \mu\nabla^2 w = 0. \tag{15}$$

where, p is pressure, (u, v) are the velocity and μ is viscosity of earth (see figure 3). Using the following boundary conditions,

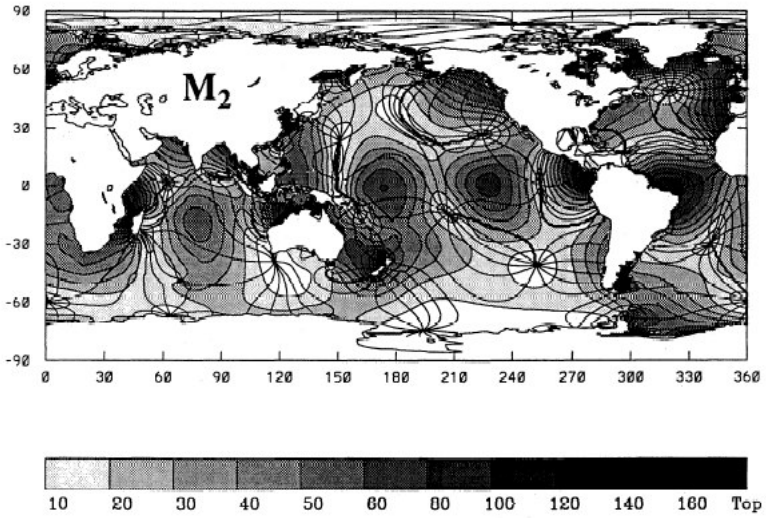


Figure 2: Cotidal map of the M_2 component. Coamplitude lines are drawn following the scaling indicated below the map. Units are in centimeters. Cophase lines are drawn with an interval of 30°, with the 0° phase as a larger drawing, referred to the passage of the astronomical forcing at Greenwich meridian [1].

$$u, w \to 0 \ as \ z \to -\infty,$$
$$\tau_{xz} = 2\mu(u_z + w_x) = 0 \ at \ z = 0,$$
$$\tau_{zz} = -p + 2\mu w_z = -\text{load} \ \equiv -\rho_w g a e^{ikx},$$

where term $\rho_w g a e^{ikx}$ gives the loading on earth surface due to ocean tide, a is the tidal amplitude and k is its horizontal wavenumber, we solve (14) and (15) with the above b.c.'s for a load of $\rho_w g a e^{ikx}$ to get an Earth surface wave displacement of

$$h\rho_w g a e^{ikx}, \tag{16}$$

where h is called "Love Number". In this case, $h = \frac{1}{2\mu k}$.

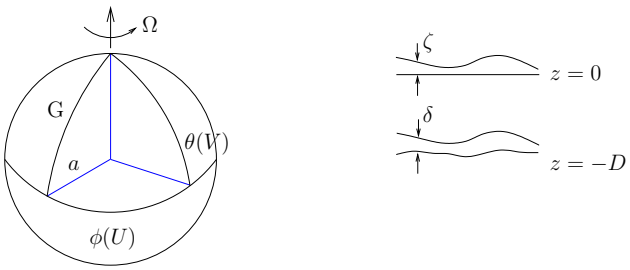


Figure 3: Solid earth tide, ζ is geocentric surface tide and δ is geocentric solid earth tide.

We can then write the tide generating potential in spherical harmonic (n) decomposition including the solid Earth tide as,

$$\Gamma_n = U_n + k_n U_n + q\alpha_n\zeta_{on} + k'_n q\alpha_n\zeta_{on}, \tag{17}$$

where U_n is obtained from astronomy, $k_n U_n$ is earth yielding to astronomical potential, $q\alpha_n\zeta_{on}$ is potential of tidal shell and $k'_n q\alpha_n\zeta_{on}$ is earth yielding to tidal potential.

Proceeding as above, solid earth tide is given as,

$$\delta_n = h_n \frac{U_n}{g} + h'_n\alpha_n\zeta_{on}. \tag{18}$$

k_n, k'_n, h_n and h'_n are all "Love numbers" similar to h in (16). From (17) and (18) we find

$$\left(\frac{\Gamma}{g} - \delta\right)_n = \left(\frac{1 + k_n - h_n}{g}\right) U_n + (1 + k'_n - h'_n)\alpha_n\zeta_{on}. \tag{19}$$

And upon summing up this series we get Farrell Green's function [2]

$$\sum_n (1 + k'_n - h'_n)\alpha_n\zeta_{on} = \int\int d\theta' d\phi' G_F(\theta', \phi'|\theta, \phi)\zeta_0(\theta', \phi'). \tag{20}$$

Now taking into account the solid Earth tide, we can rewrite Euler's equations from Lecture 2 as

$$\zeta_0 = \zeta - \delta \ , \tag{21}$$

$$u_t - (2\Omega\, sin\theta)v = -\frac{g(\zeta_0 - (\Gamma/g - \delta))_\phi}{a\, cos\theta} + \frac{F^\phi}{\rho D} \ , \tag{22}$$

$$v_t + (2\Omega\, sin\theta)u = -\frac{g(\zeta_0 - (\Gamma/g - \delta))_\theta}{a} + \frac{F^\theta}{\rho D} \ , \tag{23}$$

$$\begin{aligned}(\zeta_{ot} - \delta_t) + \frac{1}{a\, cos\theta}[(uD)_\phi + (vD\, cos\theta)_\theta] = &-\sum_n \left(\frac{1 + k_n - h_n}{g}\right) U_n \\ &- \int\int d\theta' d\phi' G_F(\theta', \phi'|\theta, \phi)\zeta_0(\theta', \phi') \ ,\end{aligned} \tag{24}$$

where U is mostly U_2, $k_2 \approx 0.29$ and $h_2 \approx 0.59$.

4 Dissipation Models

It is not easy to estimate the dissipation terms (F^θ and F^ϕ) in (22) and (23). This dissipation is mainly due to bottom drag and internal tides. If we model it as bottom drag, we get

$$\underline{F} = -\rho C_D |\underline{u}|\underline{u}, \tag{25}$$

where $C_D \approx 0.0025$ known from direct measurement in shallow water. The direct effect of (25) is that most of dissipation is limited to shallow seas where u and C_D are large. However, global tidal computations are mostly confined to deep-water zones for practical reasons (shallow water tides require much finer grid-spacing). So dissipation can only be properly represented by radiation of energy out of the model into bounding seas.

There are two main empirical models used to get an expression for dissipation in deep oceans. These are,

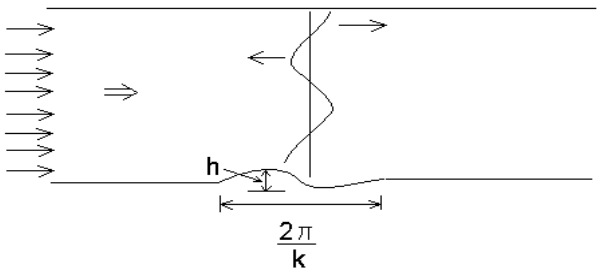


Figure 4: Jayne & Laurent model, dissipation due to barotropic tide scattering into internal tides over rough topography.

Jayne & Laurent In their runs of the Hallberg Isopycnal Model (HIM), the dissipation term is modeled as,

$$\underline{F} = -\frac{1}{2}\rho k h^2 N_b \underline{u}. \tag{26}$$

This model is based on assumption that dissipation occurs due to barotropic tide scattering into internal tides due to the rough bottom topography. h represents the height of bottom topography whose dominant horizontal wavenumber is k (see figure 4).

Arbic In his runs of the HIM for the dissipation term we have,

$$\begin{aligned} \underline{F} &= -p'_{IW} \nabla h \\ &= \rho(\nabla\chi \cdot \underline{u})\underline{\nabla} h, \end{aligned} \tag{27}$$

where,

$$\chi = \frac{N_b\sqrt{\sigma^2 - f^2}}{2\pi\sigma} \int\int \frac{h(x')}{|x - x'|} dx'dy' \tag{28}$$

for tide of frequency σ. N_b is buoyancy freqency, f is the Coriolis parameter and $h(x)$ is bottom topography. Arbic's model is based on the assumption that tidal dissipation can be calculated by finding the pressure drop in tidal currents across topographic features at the bottom.

5 Boundary Conditions

Laplace tidal equations have never been solved well enough so as to remove tides from altimetry without data assimilation. Different methods use different boundary conditions for solution of LTE at numerical coast. Some of the main boundary conditions in use are,

1. $\underline{u} \cdot \hat{h} = 0$ at the numerical coast.
   This is the most commonly used boundary condition. This boundary condition represents a no-energy-flux coast. It is important to have correct information regarding dissipation if this boundary is used since all energy must be dissipated within the system. Numerical schemes need to resolve $-\rho C_D \underline{u}|\underline{u}|$ well which requires high resolution in shallow waters.

2. $\zeta = \zeta_{obs}$ at the numerical coast.
   This boundary condition allows a energy flux ($\rho g D|\underline{u}| \cdot \hat{h}\zeta$) through the coast. The scheme is less sensitive to the details of the dissipation model used, and is less sensitive to the discretization used. However, this system can still respond resonantly. Another problem with this scheme is that observed tidal data is not easily available along all coasts.

3. $\zeta = c\ (\underline{u} \cdot \hat{h})$ at the numerical coast.
   This boundary allows energy to be dissipated at coast. In this boundary the parameter c can be adjusted so as to get results to match the observed tidal results. This leads to an energy flux of $< \rho g D c \zeta^2 >$ flowing out of the coast.

6 Energetics

If we ignore the solid Earth tide, we can derive equations of energy and perhaps estimate dissipation due to the tides as a residual. Starting with Laplace's tidal equations,

$$\begin{aligned} u_t - fv &= -g(\zeta - \Gamma/g)_x + F^x/\rho D \qquad \times \rho u D \\ v_t + fu &= -g(\zeta - \Gamma/g)_y + F^y/\rho D \qquad \times \rho v D \\ \zeta_t + (uD)_x + (vD)_y &= 0 \qquad \times \rho g \zeta \end{aligned}$$

Then multiplying by the terms at right, adding the three equations together and assuming that ρ, g and D are constant, we arrive at

$$\frac{1}{2}\rho D(u^2+v^2)_t + \frac{1}{g}\rho g(\zeta^2)_t + \nabla \cdot (\rho g \zeta u D) = \rho \zeta_t \Gamma + \nabla \cdot (\rho u D \Gamma) + u \cdot F \tag{29}$$

$$KE_t + PE_t + \nabla \cdot Eflux = \begin{array}{c} \textit{Fluid crossing} \\ \textit{equipotentials} + \\ \textit{vertically} \end{array} \begin{array}{c} \textit{Fluid crossing} \\ \textit{equipotentials} + \\ \textit{horizontally} \end{array} \begin{array}{c} \textit{Work by} \\ \textit{dissipative} \\ \textit{forces} \end{array} \tag{30}$$

where KE_t is the time derivative of kinetic energy, PE_t is the time derivative of potential energy, and $Eflux$ is energy flux.

Energy Averaged Over One Tidal Period, Integrated Over the Ocean

It is convenient to consider the energy as averaged over one tidal period. For a periodic tide, let $< \cdot >$ denote the average over one period. This will simplify the above equations, since

$$< KE_t >=< PE_t >= 0. \tag{31}$$

Then from (29) we are left with

$$\nabla \cdot < P >=< W_t > + < u \cdot F >, \tag{32}$$

where P is energy flux and W_t is the working by potential vertical and horizontal forces.

Now if we reconsider the case of the basin with no flow through its boundaries ($\vec{u} \cdot < \hat{n} >= 0$), then we further have that $\int \nabla \cdot < P > dxdy = 0$ since there can be no

net energy flux into or out of the basin. Then we can integrate the remaining terms in the energy equation over the ocean basin and find that

$$\int < W_t > dxdy = \int_{ocean} \rho\zeta_t \Gamma dxdy = -\int_{ocean} < u \cdot F > dxdy. \tag{33}$$

One caveat is that the solid Earth tide is not dissipation free, i.e. the Love numbers are complex, but this equation is true provided that they are real. Now, since $-\int < u \cdot F > dxdy$ is balanced by working of fluid moving up and down. This fluid movement ζ can be measured from global altimetry, giving an estimate if dissipation of energy due to the tides.

6.1 Including the solid earth tide

If we now include the solid Earth tide, then our third equation becomes

$$(\zeta - \delta)_t + \nabla \cdot \vec{u} D = 0 \tag{34}$$

and if we follow the same procedure as before we have

$$\begin{aligned} KE_t &= -\nabla \cdot (\rho g D \vec{u}(\zeta - \Gamma/g)) + \rho g(\zeta - \Gamma/g)\nabla \cdot \vec{u} D + u \cdot F \\ &= -\nabla \cdot (\rho g D \vec{u}(\zeta - \Gamma/g)) - \rho g(\zeta - \Gamma/g)(\zeta_t - \delta_t) + \vec{u} \cdot F \\ &= \nabla \cdot \rho D u \Gamma + \rho(\zeta - \delta)_t \Gamma - \nabla \cdot \rho g D u \zeta - \rho g \zeta(\zeta - \delta)_t + u \cdot F. \end{aligned}$$

Similarly for potential energy,

$$\begin{aligned} PE &= \int_{-D+\delta}^{\zeta} \rho g z dz = \frac{1}{2}\rho g(\zeta^2 - (-D + \delta)^2), \\ PE_t &= \rho g(\delta\delta_t - \zeta\zeta_t - D\zeta_t). \end{aligned}$$

Adding these together we find that

$$\begin{aligned} KE_t + PE_t + \nabla \cdot (\rho g u D \zeta) &= \nabla \cdot \rho D u \Gamma + \rho(\zeta - \delta)_t \Gamma + u \cdot F \\ &\quad - \rho g \zeta(\zeta_t - \delta_t) + \rho g(\zeta\zeta_t - \delta\delta_t + D\delta_t) \\ &= \nabla \cdot \rho D u \Gamma + D u \Gamma + \rho(\zeta - \delta)_t \Gamma + u \cdot F + \rho g(\zeta - \delta + D)\delta_t. \end{aligned}$$

Then if we consider the observed tide only, $\zeta_0 = \zeta - \delta$, the difference between the ocean tides and the solid earth tide, we have the energy equation.

$$KE_t + PE_t + \nabla \cdot \vec{P} = W_t + u \cdot F, \tag{35}$$

with

$$KE = \frac{1}{2}\rho D(u^2 + v^2), \tag{36}$$

$$PE = \frac{1}{2}\rho g(\zeta_0^2 + 2\zeta_0\delta + 2\delta D), \tag{37}$$

$$\vec{P} = \rho g D \vec{u}(\zeta_0 + \delta), \tag{38}$$

$$W_t = \rho\zeta_{0_t}\Gamma + \rho\nabla \cdot u D \Gamma + \rho g(\zeta_0 + D)\delta_t. \tag{39}$$

Energy Averaged Over One Tidal Period, Integrated Over the Ocean

If we again average over one tidal period, then

$$\nabla\cdot < P >=< W_t > + < u \cdot F > \tag{40}$$

Given altimeter data ζ_1 it may be possible to map $< u \cdot F >$ [3].

If we further assume that the tides are periodic as $(e^{-i\sigma t})$, then noting that in the equation for W_t that $\int_{ocean} \rho\nabla \cdot uD\Gamma = 0$ and $< \rho g D\delta_t >= 0$,

$$\int_{oc} < W_t > dxdy = \frac{1}{2} Re \left\{ int_0 - \sigma\rho\zeta_0\Gamma^* + i\sigma g\rho\zeta_0\delta^* \right\},$$

where $(\cdot)^*$ is the complex conjugate. Using the Love number decomposition from §3,

$$\begin{aligned}\int_{oc} < W_t > dxdy &= \frac{1}{2} Re \int_0 \sum_n -i\sigma\rho\zeta_{0n} \left((1+k_n)u_n^* + g\alpha_n\zeta_{0n}^* + k_n' g\alpha_n\zeta_{0n}^* \right) \\ &\qquad + i\sigma\rho\zeta_{0n}(h_n u_n^* + \delta h_n'\alpha_n\zeta_{0n}^*) \\ &= \frac{1}{2} Re \sum_n \int -i\sigma\rho(1+k_n-h_n)\zeta_{0n}u_n^*. \end{aligned} \tag{41}$$

This last equation is true provided that the solid earth tide is dissipation-less, that is to say, that the Love numbers are real. Now, since again $\int_{ocean} \nabla\cdot < P > dxdy = 0$,

$$\int_{ocean} < u \cdot F > dxdy = \int_{ocean} < W_t > dxdy = \frac{1}{2} Re \sum_n \int -i\sigma\rho(1+k_n-h_n)\zeta_{0n}u_n^* \tag{42}$$

and we can estimate the dissipation of energy due to the tides if we know the observed tides, ζ_0.

Notes by Vineet Birman and Eleanor Williams Frajka

References

[1] C. L. Provost, M. L. Genco, F. Lyard, P. Vincent, and P. Canceil, "Spectroscopy of the world ocean tides from a finite element hydrodynamic model," J. Geophys. Res. **99**, 24777 (1994).

[2] W. E. Farrell, "The effect of ocean loading on tidal gravity," EME Symposium Marees Terrestres Strasbourg: OBS. ROY. BELG. COMM 116 .

[3] G. D. Egbert and R. D. Ray, "Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data," Nature **405**, 775 (2000).

Investigator	Type of Model	Boundary Condition	Dissipation	Earth Tide
Zahel (1970)	M2, K1 tides for globe	Impermeable coast	Bottom stress in shallow water	None
Pekeris et al. (1969)	M2 tide for globe	Impermeable coast	Linearized artificial bottom stress in shallow water	None
Hendershott (1972)	M2 tide for globe	Coastal elevation specified	At coast only	None
Bogdanov et al. (1967)	M2, S2, K1, 01 tides for globe	Coastal elevation specified	At coast only	None
Tiron et al. (1967)	M2, S2, K1, 01 tides for plane model	Coastal elevation specified where data exist, otherwise impermeable coast	At coast only	None
Gohin (1971)	M2 tide for Atlantic, Indian Oceans	Coastal impedance specified	At coast only	Yielding to astronomical force only
Platzman (1972)	Normal modes for major basins, especially Atlantic	Adiabatic boundaries	Zero	-
Munk et al. (1970)	Normal mode representation of California coastal tides	Impermeable coast	Zero, although energy flux parallel to coast may reflect dissipation up coast	Yielding to astronomical force only
Cartwright (1971)	β plane representation for South Atlantic tides	Island values specified	Zero, although energy flux may reflect distant dissipation	Yielding to astronomical force only
Ueno (1964)	M2 tide for globe	Impermeable coast	None	None
Gordeev et al. (1972)	M2 tide for globe	Impermeable coast	Linearized bottom stress in shallow water	None

Table 1: Summary of Large-Scale Tidal Models.

Lecture 4: Resonance and Solutions to the LTE

Myrl Hendershott and Chris Garrett

1 Introduction

Myrl Hendershott concluded the last lecture by discussing the normal modes of the ocean, i.e. the free wave solutions of the Laplace Tidal equations. Today, we will review these concepts and extend them to so solve for specific instances of gravity and Rossby modes in simple basins. We will then discuss the computation work of Platzman [1] on the normal modes in the oceans and describe resonance in the seas.

The second half of the lecture was given by Chris Garrett and served as an introduction to his lectures during the second week. He motivates the study of tides by discussing them in the contexts of observations at several different geographical locations. He then adds to Myrl Hendershott's discussion on resonance with observations and some inferences about friction.

2 Review of Normal Modes

In plane coordinates, the Laplace Tidal equations (LTE) are given by

$$\begin{aligned} u_t - fv &= -g\zeta_x, && (1)\\ v_t + fu &= -g\zeta_y, && (2)\\ \zeta_t + D\left(u_x + v_y\right) &= 0, && (3) \end{aligned}$$

where $f = f_0 + \beta y$ and $D = D_n$, where D_n is a constant determined by an appropriate eigenvalue problem discussed in the previous lecture. To find free solutions of the LTE we look for harmonic solutions, i.e. solutions with time dependence proportional to $e^{-i\sigma t}$ which can be seen as transforming the time derivatives by

$$\frac{\partial}{\partial t} \rightarrow -i\sigma. \tag{4}$$

That is, the LTE become

$$\begin{aligned} -i\sigma u - fv &= -g\zeta_x, && (5)\\ -i\sigma v + fu &= -g\zeta_y, && (6)\\ -i\sigma\zeta + D\left(u_x + v_y\right) &= 0. && (7) \end{aligned}$$

Solving for u and ζ in terms of v, the LTE can be reduced into a single second order partial differential equation for v. We find that

$$\nabla^2 v + i\frac{\beta}{\sigma}v_x + \frac{\sigma^2 - f^2}{gD}v = 0. \tag{8}$$

If we seek plane wave solutions for v i.e. $v = e^{-i\sigma t + ikx + ily}$, then (8) implies that

$$-k^2 - l^2 - \frac{\beta k}{\sigma} + \frac{\sigma^2 - f_0^2}{gD} = 0. \tag{9}$$

This relation between k, l and σ is more useful when rewritten as

$$\left(k - \frac{\beta}{2\sigma}\right)^2 + l^2 = \left(\frac{\beta}{2\sigma}\right)^2 + \frac{\sigma^2 - f_0^2}{gD}. \tag{10}$$

We see from this equation that high and low frequency waves have significantly different character in the $k - l$ plane. High frequency waves, correspond to $\beta = 0$ which give circles centered at the origin with wavelength $\left(\left(\sigma^2 - f_0^2\right)/gD\right)^{1/2}$. These correspond to gravity modes which are characterized by $\sigma > f$ and are discussed in Section 3. For low frequencies, we have circles centered at $k = \beta/2\sigma$, $l = 0$ with radius less than $\beta/2\sigma$ so that the circles are entirely in the $k > 0$ plane. These are known as Rossby modes and are discussed in Section 4.

A third type of free wave is due to the influence of rotation. These "Kelvin" waves move along the direction of the physical boundary, e.g. the coast of a continent, and decay in magnitude as the distance from the boundary increases. For large enough basins, sum of the normal modes consist of an integral number of Kelvin waves around the boundary. An example of this will be seen in the discussion of gravity waves in a circular basin in Section 3.

3 Gravity Modes

It is instructive to solve the LTE in the context of some very simplified cases first, because the resulting motions are simplified illustrations of the true dynamics. If there is no rotation ($f = 0$), the LTE are given by

$$u_t = -g\zeta_x, \tag{11}$$

$$v_t = -g\zeta_y, \tag{12}$$

$$\zeta_t + D_0(u_x + v_y) = 0, \tag{13}$$

with the boundary condition that

$$\mathbf{u} \cdot \hat{\mathbf{n}} = 0 \tag{14}$$

at the coast, where $\hat{\mathbf{n}}$ is the unit vector normal to the coast. Substituting (11)-(12) into (13), we attain a Helmholtz equation:

$$\nabla^2 \zeta + (\sigma^2/gD)\zeta = 0, \tag{15}$$

with

$$\vec{\nabla}\zeta \cdot \hat{\mathbf{n}} = 0 \tag{16}$$

at the coast.

3.1 Square Basin

We first consider a rectangular basin (Fig. 1) with $A > B$. A general solution can be found by substituting

$$\zeta_{nm} = e^{-i\sigma_{nm}t} \cos\left(\frac{n\pi x}{A}\right) \cos\left(\frac{m\pi y}{B}\right) \tag{17}$$

into (15), which gives the dispersion relation

$$\sigma_{nm}^2 = gD_0 \left[\left(\frac{n\pi}{A}\right)^2 + \left(\frac{m\pi}{B}\right)^2\right]. \tag{18}$$

The gravest mode is then given by

$$T_{10} = \frac{2A}{\sqrt{gD_0}}, \tag{19}$$

where A is the length of the longer side. For all subsequent modes of oscillation,

$$T_{mn} > \frac{2A}{\sqrt{gD_0}}. \tag{20}$$

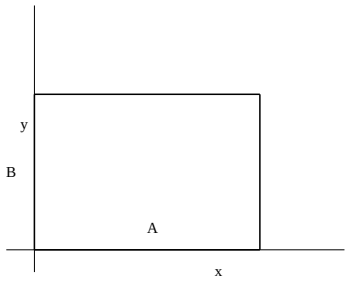


Figure 1: Rectangular ocean basin.

3.2 Circular Basin

In the real world, of course, ocean basins are not rectangular. The next general case that we can consider is that of a circular basin (figure 2). In this case, it is helpful to transform (15) to polar coordinates:

$$\zeta_{rr} + \frac{\zeta_r}{r} + \frac{\zeta_{\phi\phi}}{r^2} + \frac{\sigma^2}{gD_0}\zeta = 0, \tag{21}$$

with boundary condition

$$\zeta_r|_{r=a} = 0. \tag{22}$$

The general solution, which can be derived using separation of variables, is then given by

$$\zeta_{ns} = J_s(\kappa_{ns} r) e^{-i\sigma_{ns}t + is\phi}. \tag{23}$$

The appropriate eigenvalues for the κ_{ns} are determined by the boundary condition; that is

$$J_s'(\kappa_{ns}a) = 0. \tag{24}$$

Finally, the dispersion relation, obtained by plugging (23) into (21), is

$$\sigma_{ns}^2 = gD_0\kappa_{ns}^2. \tag{25}$$

This dispersion relation produces an ascending sequence of eigenfrequencies, σ_{ns}^2.

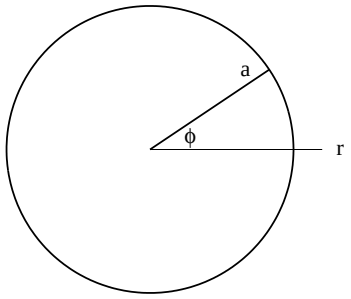


Figure 2: Circular ocean basin.

3.3 Circular Basin with Rotation

We can now add rotation to the problem in a simple way, but setting the Coriolis parameter $f = f_0$. For simplicity, we will keep the depth constant ($D = D_0$). The tidal equations are now given by

$$\begin{aligned} u_t - f_0 v &= -g\zeta_x, & (26)\\ v_t + f_0 u &= -g\zeta_y, & (27)\\ \zeta_t + D_0(u_x + v_y) &= 0, & (28) \end{aligned}$$

with the boundary condition, as before, given by (14). The Helmholtz equation can then be similarly derived, and is given by

$$\nabla^2\zeta + (\frac{\sigma^2 - f_0^2}{gD_0})\zeta = 0, \tag{29}$$

with the boundary condition

$$-i\sigma\zeta_n - f_0\zeta_s = 0 \tag{30}$$

at the coast.

Separable solutions now occur only in a circular basin. In polar coordinates, the Helmholtz equation with rotation is given by

$$\zeta_{rr} + \frac{\zeta_r}{r} + \frac{\zeta_{\phi\phi}}{r^2} + \frac{\sigma^2 - f_0^2}{gD_0}\zeta = 0, \tag{31}$$

with boundary condition

$$-i\sigma\zeta_r - \frac{f_0\zeta_\phi}{a} = 0. \tag{32}$$

The general solution is again given by (23), but now with a modified dispersion relation,

$$\sigma_{ns}^2 = f_0^2 + gD_0\kappa_{ns}^2. \tag{33}$$

For $s = 0$ (that is, no wave in the radial direction), the solution, boundary condition, and dispersion relation are given by

$$\begin{aligned} \zeta_{n0} &= J_0(\kappa_{n0}r)e^{-i\sigma_{0n}t}, & (34)\\ J_0'(\kappa_{n0}a) &= 0, & (35)\\ \sigma_{n0}^2 &= gD_0\kappa_{n0}^2. & (36) \end{aligned}$$

The boundary condition given by (35) again fixes κ_{n0}, with an ascending sequence of positive eigenfrequencies following. This case is just like the case of no rotation ($f_0 = 0$), except that particle paths will now no longer be radial. For $s \neq 0$, substituting the dispersion relation given by (33) into (32) gives

$$-i\sigma\kappa_{ns}J_s'(\kappa_{ns}a) = \frac{f_0 is}{a}J_s(\kappa_{ns}a) = 0, \tag{37}$$

which fixes κ_{ns}. Solving (37) for σ and substituting (33), we have

$$\frac{sJ_s(\kappa a)}{\kappa a J_s'(\kappa a)} = \pm\sqrt{1 + \frac{(\kappa a)^2}{\beta}}, \tag{38}$$

where

$$\beta := \frac{f_0^2 a^2}{gD_0} \tag{39}$$

is called the Lamb Parameter.

Solutions admitted by this problem are given by the roots of (38), as illustrated in figure 4. For a given s, and $f_0 \to 0$ (and, consequently, $\beta \to 0$) the roots are given by $(\kappa a)^2 > 0$ pairs near the $J_s'(\kappa a) = 0$ line. As f_0 and β increase, $(\kappa a)^2$ pairs move farther apart, and a new root appears, with $(\kappa a)^2 < 0$. For this case, κ, and therefore the argument of the Bessel function in (23) will be an imaginary number. The radial part of the solution is then given by a modified Bessel function,

$$I_s(\kappa r) = i^{-n}J_n(ix) = e^{-n\pi i/2}J_n(xe^{i\pi/2}), \tag{40}$$

which decays away from the boundary. Figure 3 shows the dispersion relation given in (33) for a fixed s and different n. The pairs of eigenfrequencies corresponding to the κa pairs in figure 4 can be readily seen; they are gravity waves propagating in opposite directions. The $(\kappa a)^2 < 0$ mode is the Kelvin Wave.

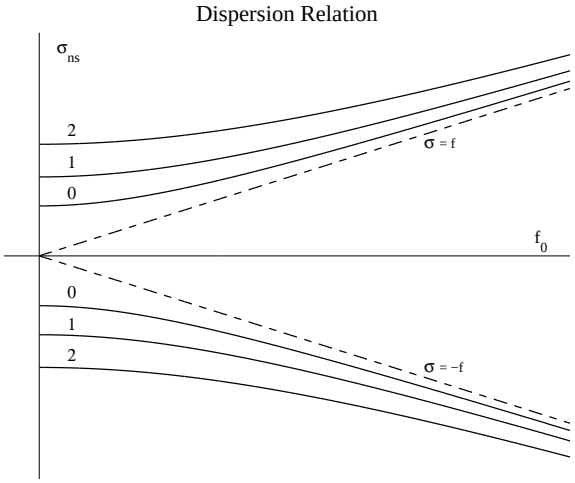


Figure 3: The dispersion relation (33) for different values of n. The Kelvin modes are shown by the dashed line.

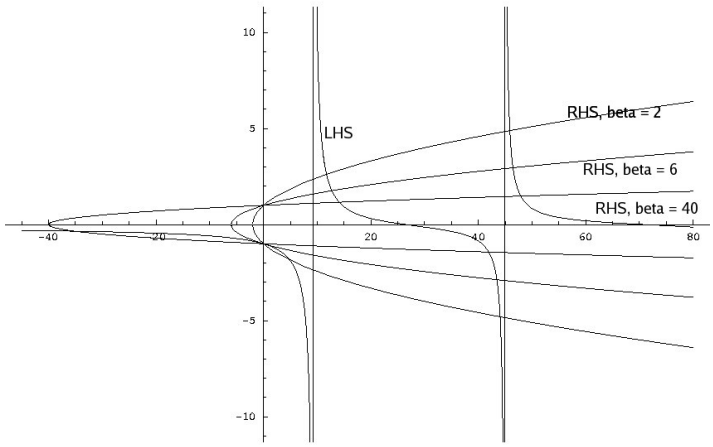


Figure 4: Illustration of the roots of (38). The curves opening to the right correspond to the right hand side of (38) for different β (with β increasing for tighter curves. The asymptotic function is the left hand side of (38). For this case, $s = 2$.

4 Rossby Modes

4.1 Linear Models for Rossby Waves

We now return to the Rossby modes. As discussed earlier, these are the low frequency solutions of the LTE in plane coordinates. Now defining the vorticity, ξ, in the usual way, i.e.

$$\xi = v_x - u_y, \tag{41}$$

we rewrite the shallow water wave equations to give

$$\frac{D}{Dt}\left(\frac{\xi + f}{D}\right) = 0, \tag{42}$$

where D/Dt is the material derivative, i.e.

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}.$$

Recalling that $f = f_0 + \beta y$ and $D = D_0 + \zeta$ we linearize (42) with respect to ξ, ζ, u and v to find an equation for the time evolution of the vorticity.

$$\begin{aligned}\frac{D}{Dt}\left(\frac{\xi + f}{D}\right) &= \frac{D}{Dt}\left(\frac{\xi + f_0 + \beta y}{D_0} - \frac{(f_0 + \beta y)\,\zeta}{D_0^2}\right) \\ &= \frac{\xi_t}{D_0} - \frac{f_0\zeta_t}{D_0^2} + \frac{\beta v}{D_0},\end{aligned} \tag{43}$$

where we have assumed that $\beta y \ll f_0$ to neglect the term proportional to $\beta y\zeta$.

Since Rossby waves are low frequency waves, $u_t \ll f_0 v$ and $v_t \ll f_0 u$ so that

$$-f_0 v \approx -g\zeta_x \quad f_0 u \approx -g\zeta_y. \tag{44}$$

Substituting (44) into (43), we find that we can rewrite (43)

$$\frac{g}{f_0 D_0}\left(\zeta_{xx} + \zeta_{yy}\right) - \frac{f_0}{D_0^2}\zeta_t + \frac{\beta g}{f_0 D_0}\zeta_x = 0. \tag{45}$$

Multiplying (45) through by $f_0 D_0/g$ (45) becomes

$$\nabla^2\zeta_t - \left(\frac{f_0^2}{gD_0}\right)\zeta_t + \beta\zeta_x = 0. \tag{46}$$

By considering the boundary condition on the coast, we can further simplify the vorticity equation for Rossby waves. As in the earlier discussion of gravity waves, at the coast we must have

$$\mathbf{u}\cdot\hat{\mathbf{n}} = 0,$$

so that if we define a local coordinate system at each point on the coast with the $\hat{\mathbf{s}}$ direction tangent to the coast, we must have $\zeta_s = 0$ so that

$$\zeta_{coast} = \Gamma(t), \tag{47}$$

i.e. that value of ζ is the same everywhere on the coast though it may be allowed to vary in time. For sufficiently short waves,

$$\nabla^2 \zeta_t \gg \left(\frac{f_0^2}{g D_0} \zeta_t \right), \tag{48}$$

so that to leading order (46) becomes

$$\nabla^2 \zeta_t + \beta \zeta_x = 0. \tag{49}$$

Since the boundary condition is independent of space, the time dependent boundary condition can, to leading order, be absorbed into ζ, so that (49) can be solved under the condition

$$\zeta_{coast} = 0. \tag{50}$$

4.2 Rossby Waves in a Square Basin

Suppose that ζ is periodic in time with frequency σ. That is, suppose that ζ can be written in the form

$$\zeta(x, y, t) = \Re \left\{ e^{-i\sigma t} \Phi(x, y) \right\}. \tag{51}$$

Substituting this expression for ζ into (49), we find that Φ satisfies the following boundary value problem.

$$\nabla^2 \Phi + \frac{i\beta}{\sigma} \frac{\partial \Phi}{\partial x} = 0, \quad \Phi_{coast} = 0. \tag{52}$$

To remove the x-derivative from (52) we further substitute

$$\Phi = e^{-i\beta x / 2\sigma} \phi(x, y) \tag{53}$$

into equation 52 we find that ϕ satisfies the boundary value problem

$$\nabla^2 \phi + \lambda^2 \phi = 0, \phi_{coast} = 0, \tag{54}$$

where $\lambda^2 = \beta^2 / 4\sigma^2$.

Consider Rossby waves in the rectangular basin $0 \leq x \leq x_0$, $0 \leq y \leq y_0$. Using separation of variables, we write

$$\phi(x, y) = X(x) Y(y) \tag{55}$$

and see that X and Y both satisfy equations for the harmonic oscillator. Therefore the boundary conditions imply that any function of the form

$$\phi = \phi_{mn} = \sin \frac{m\pi x}{x_0} \sin \frac{n\pi y}{y_0}, \tag{56}$$

with m, n positive integers, satisfies the boundary value problem and that σ is the corresponding eigenvalue,

$$\sigma = \sigma_{mn} = -\frac{\beta}{2\pi} \frac{1}{\left[\left(m^2 / x_0^2 \right) + \left(n^2 / y_0^2 \right) \right]^{1/2}}. \tag{57}$$

Note that the highest frequency modes are the ones with the smallest values of m and n. Substituting, our solution for ϕ_{nn} into (53) and subsequently into (51), we find that for each normal mode

$$\zeta = \cos\left[\frac{\beta x}{2\sigma_{mn}} + \sigma_{mn} t\right] \sin\left(m\pi \frac{x}{x_0}\right) \sin\left(n\pi \frac{y}{y_0}\right). \tag{58}$$

This solution consists of a carrier wave moving to the left modulated by sine function serving to satisfy the boundary conditions. The sine functions also create stationary nodes while the cosine function creates nodes which move to the left.

4.3 Circular Basin

Now suppose we have a circular basin with radius a. The definitions for ϕ follow exactly the same as for the square basin leaving the boundary value problem

$$\nabla^2 \phi + \lambda^2 \phi = 0, \quad \phi\,(r = a) = 0, \tag{59}$$

with $\lambda^2 = \beta^2/4\sigma^2$.

We again use separation of variables to write

$$\phi = R\,(r)\,\Theta\,(\theta) \tag{60}$$

to find that Θ and R satisfy the harmonic oscillator and Bessel Equations respectively. Therefore, the boundary conditions and single-valuedness imply that any function of the form

$$\phi = \phi_{nm} = \cos\,(m\theta + \alpha)\, J_m\,(k_{nm} r)\,, \tag{61}$$

with α an arbitrary phase angle, satisfies the boundary conditions under the conditions that m is a positive integer, J_m is the Bessel Function of order m, and

$$\xi_{mn} = k_{nm} a \tag{62}$$

is the n^{th} zero of the Bessel function of order m. The eigenvalue for σ is then

$$\sigma_{mn} = -\frac{\beta}{k_{nm}^2 a^2}, \tag{63}$$

and the corresponding eigenfunction is

$$\zeta_{mn} = \cos\left[\frac{\beta x}{2\sigma_{mn}} + \sigma_{mn} t\right] \cos\,(m\theta + \alpha)\, J_m\,(k_{nm} r)\,. \tag{64}$$

Therefore these eigenfunctions also represent a carrier wave moving to the left creating moving nodes modulated by an envelope of functions which create stationary nodes. In this case the stationary modes are radial lines corresponding to zeros of $\cos\,(m\theta + \alpha)$ and circles corresponding to zeros of $J_m\,(k_{nm} r)$.

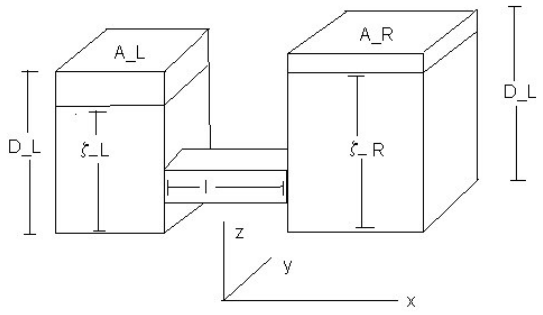


Figure 5: The Helmholtz Oscillator

5 Helmholz "Mass Exchange" Mode

We can solve for an additional normal mode, where fluid moves between two basins. The problem is illustrated in Fig. 5 in which A_L, D_L and A_R, D_R represent the cross-sectional area and depth of the left and right basins respectively. We further take l, w, and d to be the length, width, and depth of the connecting channel.

If we take $D_L, D_R \gg d$ so that the fluid height, ζ, in each basin can be taken as virtually constant within the basin and the connecting channel is always full. Define ζ_L and ζ_R as the time dependent fluid heights in the left and right basins, respectively. The total fluid volume in the system is

$$V = A_L\zeta_L + A_R\zeta_R + lwd \tag{65}$$

and must be conserved conserved. Differentiating (65) with respect to time, we derive the volume conservation equation:

$$A_L\left(\zeta_L\right)_t + A_R\left(\zeta_R\right)_t = 0. \tag{66}$$

Since we have taken ζ_L and ζ_R to be independent of space, we can express the time rate of change of the volume flux into the right basin as

$$\frac{d}{dt}A_R\left(\zeta_R\right)_t = A_R\left(\zeta_R\right)_{tt}.$$

Any fluid that enters the right basin must enter through the connecting channel. If we assume the velocity of the fluid in the channel is entirely in the lengthwise direction, then the fluid flux into the right basin is uwd, where u is the fluid velocity in the lengthwise direction and taken to be parallel to the x-axis. Therefore, equating the time rate of change of fluid into the right basin gives us

$$A_R\left(\zeta_R\right)_{tt} = u_t wd. \tag{67}$$

Since we have assumed that the flow in the channel is uniform and unidirectional, the momentum equation in the channel is $u_t = -g\zeta_x$. At the left edge of the channel, $\zeta = \zeta_L$

while $\zeta = \zeta_R$ at the right edge. Therefore, since the length of the channel is l, $\zeta_x \approx (\zeta_R - \zeta_L)/l$. Therefore, we can rewrite (67) as

$$A_R (\zeta_R)_t \, t \approx -wdg \frac{\zeta_R - \zeta_L}{l}. \tag{68}$$

Solving (65) for ζ_L and substituting into (68) we find that ζ_R satisfies the equation for a simple harmonic oscillator. That is

$$(\zeta_R)_{tt} + gwd \left(\frac{1}{A_R} + \frac{1}{A_L} \right) \zeta_R = K, \tag{69}$$

where K is a constant. This equation implies that the Helmholtz mass exchange mode does represent a simple wave solution of the shallow water wave equations with

$$\omega^2_{Helmholtz} = gwd \left(\frac{1}{A_R} + \frac{1}{A_L} \right). \tag{70}$$

6 Platzman's Analysis

Platzman et al. [1] computed approximate normal modes of the world oceans by computing the normal modes in the 8 to 80-hour spectrum of a numerical model of the LTE. Platzman's analysis focuses on large-scale features of the direct response of the deep ocean to the tidal potential, rather than coastal tides. The discretization of the model makes it an eigenvalue problem with 2042 total degrees of freedom. The eigenvectors have the form

$$(\xi, \phi, \psi) = Re[(Z, \Phi, \Psi) e^{i\sigma t}], \tag{71}$$

where σ is the eigenfrequency, ξ is the free surface elevation, ϕ is velocity potential, and ψ is the volume streamfunction.

Platzman et al. visualized the different normal mode motions described in the previous sections, by contouring the amplitude and phase of the mean elevation, as well as the energy density and flux, over the model grid. It is instructive to examine some example maps, which illustrate different normal modes.

Figure 6 shows a sample analysis from Platzman et al., of a topographic vorticity wave, as contained by the 14th computed normal mode, which has a period of $T = 33h$. The height contour plot in this figure (left panel) shows a sort of height "dome", and the phase lines (perpendicular to the lines of height) show the counterclockwise rotation about an amphidrome where the phase contours are anchored. The energy diagram (right panel) shows that energy is transported in an anticyclonic gyre, with the energy flow largely parallel to the phase velocity of the wave. (Since the study only looked for modes with periods between 8 and 80 hours, it couldn't clearly resolve any planetary vorticity waves.)

Similarly, Mode 12 (Fig. 7) can be understood, given the height and energy contours as the Helmholtz resonator mode described in section 5. The energy diagram for this mode shows a nearly uniform phase across each ocean, and the elevation contours could perhaps be interpreted as having a "node-like band" at the junction between the two oceans. A clearer argument for the Helmholtz resonator analogy is given by Fig. 8, which simply

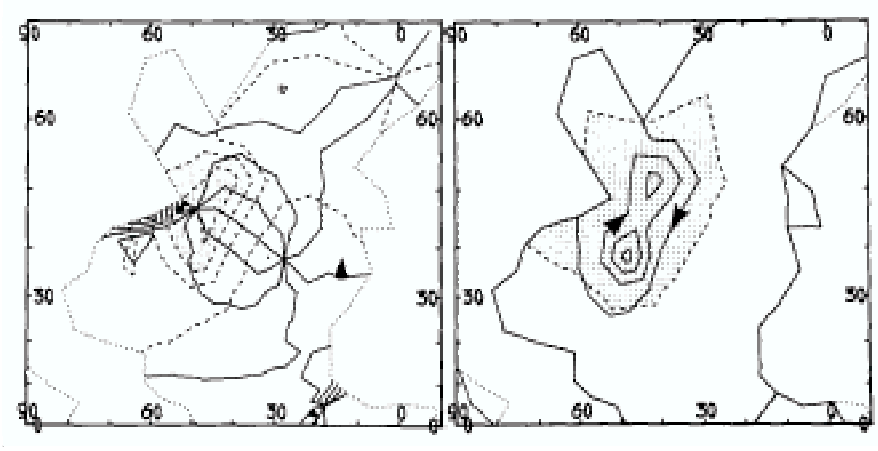


Figure 6: Platzman et al.'s computed mode 14, a vorticity wave near Newfoundland. Elevation is contoured in the left panel, total energy density on the right. Arrows are placed on the contour of zero phase, and point in the direction of phase propagation.

shows a polar plot of the amplitude and phase of the fluctuations in regional volume which are induced by that particular mode.

This figure suggests that the volume of the Arctic and North Atlantic oceans are in balance with the Indian and South Atlantic Oceans.

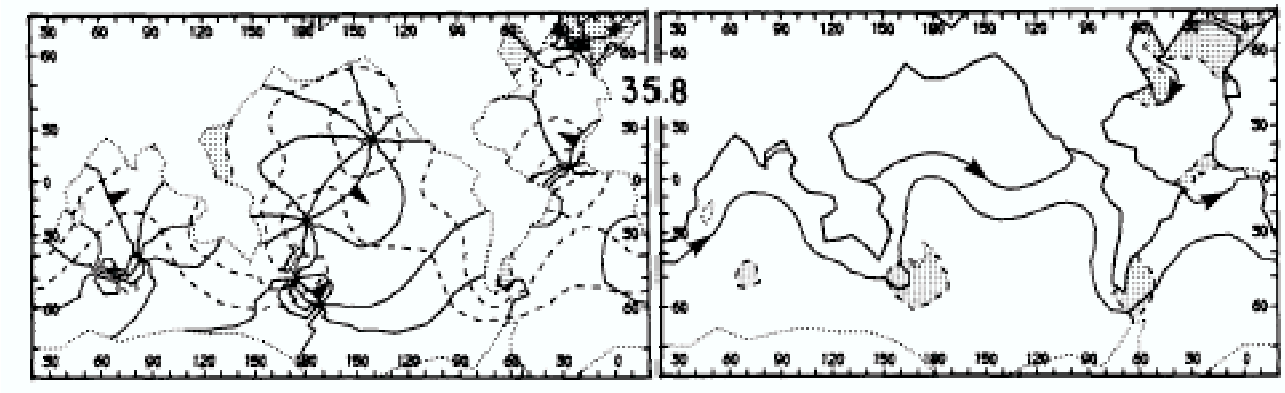


Figure 7: Platzman et al.'s computed mode 12, which can be interpreted as a Helmholtz resonator mode.

Mode 12 carries less than 30% rotational kinetic energy. Mode 15 (Fig. 9) also carries little rotational energy, though it also has a clear Kelvin Wave component in the southern ocean, with phase lines parallel to the coast of Antarctica, and eastward energy propagation.

As for gravity modes (Section 3), Platzman et al.'s model resolves 5 slow ($T < 24h$) gravity modes, but these tend to take on the appearance of vortical modes in the presence

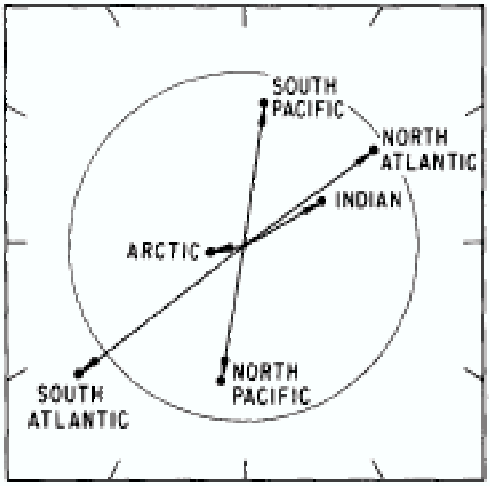


Figure 8: Volume-fluctuation vectors for Platzman et al.'s mode 12.

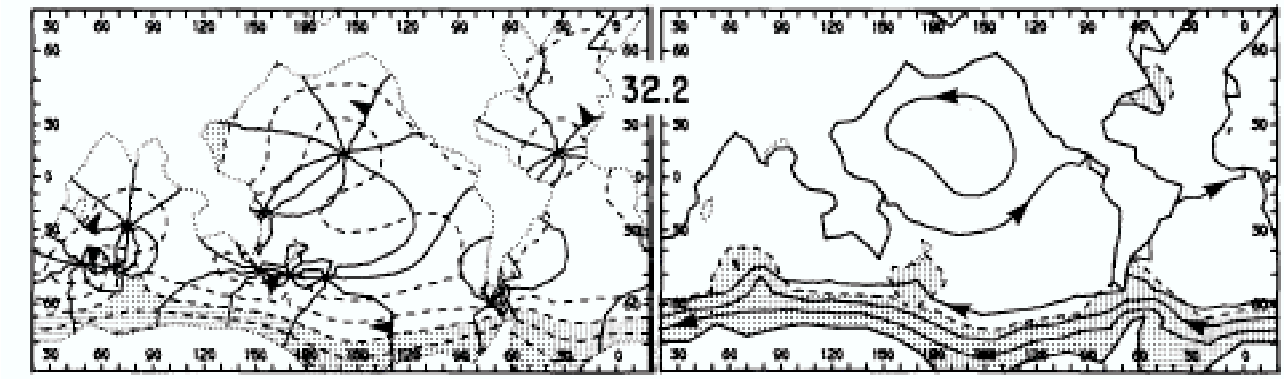


Figure 9: Platzman et al.'s computed mode 15, which contains a Kelvin Wave component.

of bottom topography. An example of a fast gravity wave is shown in figure 10, of mode 28. In this example, the arctic ocean seems to be strongly excited, and the arctic energy density is much larger than the global average.

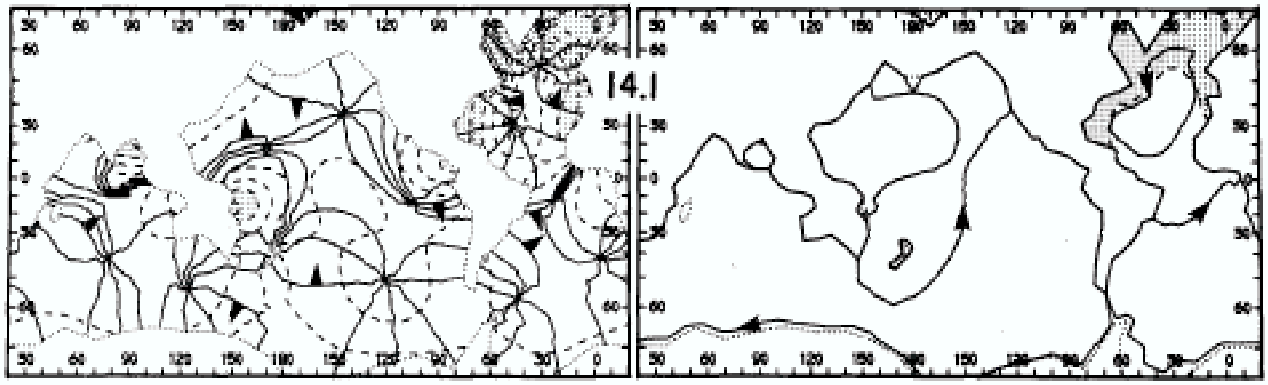


Figure 10: Platzman et al.'s computed mode 28, a gravity wave.

7 Resonance

The response of the oceans to tidal forcing at frequency ω, given a spectrum of ocean normal modes, can be written as

$$\zeta(\mathbf{x},t) = \mathbf{Re}\sum_n \frac{B_n(\omega)}{\omega_n - \omega - \frac{1}{2}iQ_n^{-1}\omega_n} S_n(\mathbf{x})e^{-i\omega t}, \tag{72}$$

where ω_n is the frequency of the nth normal mode of the oceans, $S_n(\mathbf{x})$ is the corresponding eigenfunction, and $B_n(\omega)$ is a complex factor that needs to be somehow determined. Q_n is a dissipation factor; that is, the nth normal mode dissipates a fraction $2\pi Q_n^{-1}$ of its energy per cycle. The response function given in equation (72) assumes that the functions $S_n(\mathbf{x})$ form a complete set. Equation (72) also makes the implicit assumption that the normal modes are linear waves. This assumption does not really hold in shallow water, where the nonlinear friction terms become significant. It is instructive to examine equation (72) for a hypothetical case where one mode (say,the zeroth mode) dominates, and its corresponding spatial response is constant, such that $B_0(\omega)S_0(\mathbf{x}) = C_0$. Then we can write

$$Ae^{i\Theta} = \frac{C_0}{\omega_0 - \omega - \frac{1}{2}iQ^{-1}\omega_0}. \tag{73}$$

This response function is plotted in figure 11. The age of the tide (i.e. the time lag between the tidal potential and the maximum of the response) can be found from $d\Theta/d\omega$, which, in this simple case, is given by

$$\frac{d\Theta}{d\omega} = \frac{\frac{1}{2}Q_0^{-1}\omega_0}{(\omega_0 - \omega)^2 + (\frac{1}{2}Q_0^{-1}\omega_0)^2}. \tag{74}$$

For the single-mode case, the maximum value is $2Q_0/\omega_0$ (at $\omega = \omega_0$). If the tide were dominated by single ocean mode, the age of the tide would be positive, and the same everywhere. Since

$$Q_0 \geq \frac{1}{2}\omega_0 \frac{d\Theta}{d\omega} \tag{75}$$

the age of the tide would put a lower limit on the quality factor Q_0.

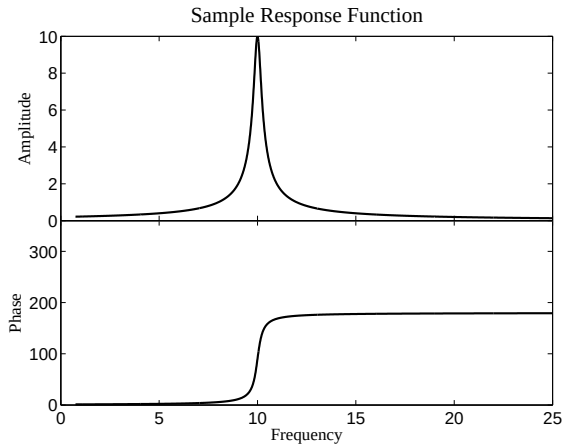


Figure 11: Sample response function for a single normal mode, where we have chosen $w_0 = 1$, $Q_0 = 25$, and $C_0 = 1$.

8 Comments on Inferences drawn from Ocean Models

Looking at the models for the ocean's normal modes we have just discussed, we can draw some interesting conclusions. Taking the dynamical solution of the "time scale" problem for the lunar orbit evolution as being solved by Hansen [2], Webb [3], Ooe et al. [4], and Kagan and Maslova [5], the models generate significantly smaller torques in the past than that implied by earth's present rotation rate. Platzman's work explains that the ocean has many different normal modes, some of which have frequencies close to the main tidal spectral lines. The modes grow more complex as frequency increases. In the past, the Earth's faster rotation forced tidal frequencies higher so that the modes with near-tidal frequencies were less well matched spatially to the large-scale tidal forcing, so that they were less easily excited. Therefore the torques were smaller since the tidal admittance to the tidal potential was also decreased from its current day value.

9 Why Ocean Tides are Back in Fashion

The LTE were written down 1776, so why are we still studying tides in 2004? There are several reasons: understanding why certain places see such large tides, understanding tidal dissipation from internal tide mixing, and the possibility of using the tides as a renewable energy source. We will discuss the first of these in this lecture.

Name	Type	Period(h)	Amp.
Q_1	L	26.87	0.0641
O_1	L	25.82	0.3800
P_1	S	24.07	0.2011
K_1	L/S	23.93	0.6392
N_2	L	12.66	0.0880
M_2	L	12.42	0.3774
S_2	S	12.00	0.1089

Table 1: Dominant constituents at Victora

10 The Tides at Victoria

If the moon were perfectly aligned with the equator, its motion would only cause semi diurnal tides. However, the moon orbits the earth at a declination angle which means that it also forces a diurnal tide. The dominant semi diurnal and diurnal tides at Victoria, BC, can be seen in Table 1.

The interactions of these tides account for the daily variations in the tides. M_2 and S_2 beat over the spring/neap cycle. K_1 and O_1 beat over a lunar month to allow for changes in the moon's declination which modulates the lunar diurnal tide. K_1 and P_1 beat over a year to allow for changes in the sun's declination which accounts for changes in the solar diurnal tide. Furthermore, M_2 and N_2 beat to provide a correction allowing for ellipticity in the moon's orbit which also effects the lunar diurnal tide.

Since K_1 has the largest amplitude we expect the tides in Victoria to look mostly diurnal. That is there should be approximately one maximum and one minimum tide during the day with some variations due to the other tides. However, in the third week of March 2004, for example, there were several consecutive days where there are two distinct maxima and minima. This is an example of the beating discussed in the previous paragraph. At the spring perigee, the sun and the moon were very close to the equatorial plane which diminishes the diurnals allowing the semidiurnals to appear more prominently, particularly at spring tides.

11 The World's Highest Tides

Two Canadian locations have recently both claimed that they have the world's largest tides. In the upper part of the Bay of Fundy, the tides rise up to 17m. Ungava Bay in Quebec also has tides that can reach at least 16.8m. Both of these maxima occur in an 18.61 year nodal cycle.

There have been several "explanations" offered by various sources for why the tides in the Bay of Fundy are so large. These include the idea that the Bay of Fundy is the closest point to the moon at these times and the belief that large tidal currents from the Indian Ocean can enter the bay because it opens to the south! In actuality, these large tides are more likely connected to a resonance phenomenon [6], which will be discussed in the next section.

12 Friction and Resonance

The time lag between the spring/neap cycles of the different tidal potential constituents, and the actual observed response of the oceans, is called the age of the tide [7]. It can be explained mathematically in terms of the response function given in (72) introduced in Section 7. The effect of friction can be examined more closely by approximating the energy dissipated due to friction, which is proportional to $u|u|$. If we assume that the tidal current u is the sum of a primary tidal constituent and a secondary, weaker one,

$$u = u_0(\cos\omega_1 t + \epsilon\cos\omega_2 t), \tag{76}$$

(where ϵ is a small parameter) then it can be shown that

$$u|u| = \frac{8}{3\pi}u_0^2(\cos\omega_1 t + \frac{3}{2}\cos\omega_2 t + ...), \tag{77}$$

where we have neglected higher-order harmonics. The 3/2 coefficient in the second term of (77) shows that the weaker tide feels a stronger frictional effect. In other words, the Q factor for the weaker tide will be 2/3 that of the stronger tide.

The response of the Bay of Fundy and the Gulf of Maine implies a resonant period of about 13.3 hours and a Q, for M_2, of about 5. On the west coast, the tidal response of the Strait of Georgia and Puget Sound can be modeled as a Helmholtz Resonance, where the tidal current enters a bay of surface area A through a channel of length L and cross section E. We will assume that the bay is small enough such that the height of the water surface rises uniformly everywhere. If the tide outside the bay is $a\cos\omega t$, the level inside is $\mathbf{Re}(a'e^{-i\omega t})$, where a' is the response amplitude in the bay. The current in the entrance channel is $\mathbf{Re}(ue^{-i\omega t})$. The continuity equation is is then

$$-i\omega a'A = Eu. \tag{78}$$

The channel dynamics are a balance between acceleration, friction, and the pressure gradient. The momentum equation can be written as

$$-i\omega u + g(a' - a)/L = -\lambda u, \tag{79}$$

where λ is simply a linear damping factor. As shown in equation (77), λ will be 50% larger for the constituents with weaker currents than the dominant M_2. The response in the bay will then be given by

$$a' = \frac{a}{1 - \frac{w^2}{w_0^2} - \frac{i\omega\lambda}{w_0^2}}, \tag{80}$$

where $\omega_0 = (\frac{gE}{AL})^{1/2}$ is the resonant frequency.

The frequency dependent response curve for a Helmholtz resonator given in equation (80). The response for a rectangular bay (not shown) can also be fitted to the data, that is, the tidal constituent frequencies and the corresponding amplitude and phase lag measurements.

For both cases, it seems that the observed tides straddle a resonant frequency ω_0, with $Q = \omega_0/\lambda$. For both cases, the fit returns a $Q \simeq 2$, which is low (and thus corresponds to high friction). This seems to be consistent with the high friction required by numerical models.

13 Nodal Modulation

The effect of the nodal modulation of the tide (due to the variation of the moon's declination) can be similarly written as a response curve, proportional to

$$a \propto \frac{a_0(1+\epsilon)}{1-\frac{\omega}{\omega_0}+\frac{1}{2}iQ_0^{-1}\frac{a}{a_0}}, \tag{81}$$

where the $1+\epsilon$ factor includes the effect of the modulation of semidiurnal forcing and the multiplier in the friction term allows for quadratic friction [8].

Away from resonance, the friction term becomes insignificant, and $a \propto 1+\epsilon$. Near resonance, however, the friction term becomes important, and $a \propto 1+\frac{1}{2}\epsilon$. Indeed, the actual modulation in the Bay of Fundy is less than in other areas.

Notes by David Vener and Lisa Neef

References

[1] G. Platzman, G. Curtis, K. Hansen, and R. Slater, "Normal modes of the world ocean. part ii: Description of modes in the period range 8 to 80 hours," J. Phys. Ocean. **11**, 579 (1981).

[2] K. Hansen, "Secular effects of oceanic tidal dissipation on the moon's orbit and the earth's rotation," Review of Geophysics **20**, 457 (1982).

[3] D. Webb, "Tides and the evolution of the earth moon system," Geophys. J. Roy. Astr. Soc. **70**, 261 (1982).

[4] H. S. M. Ooe and H. Kinoshita, "Effects of the tidal dissipation on the moon's orbit and the earth's ratation," Geophysical Monograph **59**, 51 (1990).

[5] B. Kagan and N. Maslova, "A stochastic model of the earth-moon tidal evolution accounting for the cyclic variations of resonant properties of the ocean: An asymptotic solution," Earth, Moon and Planets **66**, 173 (1994).

[6] C. Garrett, "Tidal resonance in the bay of fundy," Nature **238**, 441 (1972).

[7] C. Garrett and W. Munk, "The age of the tide and the 'Q' of the oceans.," Deep-Sea Res. **18**, 493 (1971).

[8] L. Ku, D. Greenberg, C. Garrett, and F. Dobson, "The nodal modulation of the m_2 tide in the bay of fundy and gulf of maine," Science **230**, 69 (1985).

Lecture 5: The Spectrum of Free Waves Possible along Coasts

Myrl Hendershott

1 Introduction

In this lecture we discuss the spectrum of free waves that are possible along the straight coast in shallow water approximation.

2 Linearized shallow water equations

The linearized shallow water equations(LSW) are

$$u_t - fv = -g\zeta_x, \tag{1}$$
$$v_t + fu = -g\zeta_y, \tag{2}$$
$$\zeta_t + (uD)_x + (vD)_y = 0. \tag{3}$$

where D, η are depth and free surface elevations respectively.

3 Kelvin Waves

Kelvin waves require the support of a lateral boundary and it occurs in the ocean where it can travel along coastlines.

Consider the case when $u = 0$ in the linearized shallow water equations with constant depth. In this case we have that the Coriolis force in the offshore direction is balanced by the pressure gradient towards the coast and the acceleration in the Longshore direction is gravitational. Substituting $u = 0$ in equation (1) through (3) (assuming constant depth D) gives

$$fv = g\zeta_x, \tag{4}$$
$$v_t = -g\zeta_y, \tag{5}$$
$$\zeta_t + Dv_y = 0, \tag{6}$$

eliminating ζ between equation (4) through equation (6) gives

$$v_{tt} = (gD)v_{yy}. \tag{7}$$

The general wave solution can be written as

$$v = V_1(x, y + ct) + V_2(x, y - ct), \tag{8}$$

where $c^2 = gD$ and V_1, V_2 are arbitrary functions. Now note that ζ satisfies

$$\zeta_{tt} = (gD)\zeta_{yy}. \tag{9}$$

Hence we can try $\zeta = AV_1(x, y + ct) + BV_2(x, y - ct)$. From equation (4) or equation (5) we get $A = -\sqrt{H/g}$, $b = \sqrt{H/g}$. V_1 and V_2 can be determined as follows. From equation (4) we have

$$\frac{\partial v}{\partial x} = -\frac{f}{\sqrt{gD}}V_1, \tag{10}$$

$$\frac{\partial v}{\partial x} = \frac{f}{\sqrt{gD}}V_2. \tag{11}$$

Solving the two equations we get

$$V_1 = V_{10}(0, y + ct)e^{-\frac{x}{R}}, \tag{12}$$

$$V_2 = V_{20}(0, y - ct)e^{\frac{x}{R}}, \tag{13}$$

where $R = \sqrt{gD/f}$. Since the second solution increases exponentially from the boundary it is deemed physically unfit and so the general solution can be written as

$$u = 0, \tag{14}$$

$$v = F(y + ct)e^{-\frac{x}{R}}, \tag{15}$$

$$\zeta = -\sqrt{\frac{D}{g}}F(y + ct)e^{-\frac{x}{R}}, \tag{16}$$

where F is some arbitrary function. Since we have exponential decay away from the boundary the kelvin wave is said to be trapped.

4 Poincare Continuum

Let us relax the condition $u = 0$ and keep the equations (1) through (2). With constant f and a flat bottom we can seek a Fourier solution,with u, v, ζ taken as constant factor times the function $e^{(-i\sigma t+ilx+iky)}$. The system of equations (1),(2),(3) becomes

$$-i\sigma u - fv = -igl\zeta, \tag{17}$$

$$-i\sigma v + fu = -igk\zeta, \tag{18}$$

$$-i\sigma\zeta + D(ilu + ikv) = 0. \tag{19}$$

Eliminating u, v from equations (17) through (19) we get the following dispersion relation,

$$\sigma^2 = f^2 + gD(l^2 + k^2). \tag{20}$$

This implies $\sigma^2 > f^2 + gDk^2$. Hence inside the hyperbola $\sigma^2 = f^2 + gDk^2$ there is a continuum of waves with a given frequency σ and this region is called as Poincare's continuum.

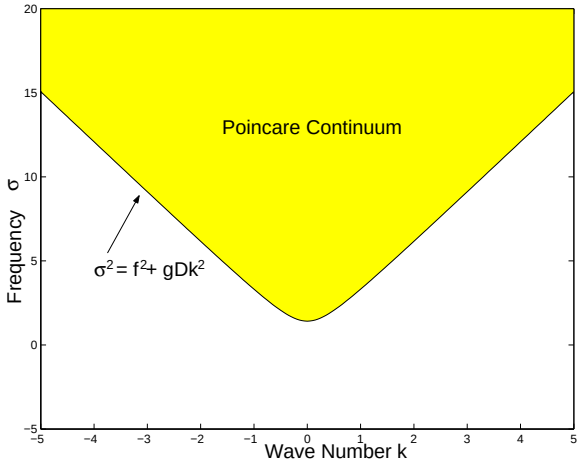


Figure 1: Poincare continuum.

5 Reflection along a straight coast

Consider the reflection of a wave $ae^{(-i\sigma t-ilx+iky)}$ at a straight coast(figure 2). Assuming constant depth we now derive the equation for the reflected wave. At the coast we have the condition that the horizontal velocity $u = 0$. The reflected wave will have the form $ae^{(-i\sigma t+ilx+iky)} + ae^{(-i\sigma t-ilx+iky)}$. Substituting in equations (1),(2) we get,

$$-i\sigma u - fv = -\zeta_x, \tag{21}$$
$$-i\sigma v + fu = -\zeta_y. \tag{22}$$

Solving for u we get,

$$u = \frac{-i\sigma g\zeta_x + fg\zeta_y}{\sigma^2 - f^2}. \tag{23}$$

Hence the boundary condition $u = 0$ implies that

$$-i\sigma g\zeta_x + fg\zeta_y = 0. \tag{24}$$

Using (24) we get,

$$b = a\frac{-l\sigma + ifk}{l\sigma + ifk}. \tag{25}$$

Hence the reflected waves can be represented as

$$a\left[e^{(-i\sigma t+ilx+iky)} + \frac{-l\sigma + ifk}{l\sigma + ifk}e^{(-i\sigma t+ilx+iky)}\right]. \tag{26}$$

6 Waves on a beach with a sloping bottom

We now consider the spectrum of free waves possible in a beach with a sloping bottom where the depth is given by $D = ax$. Let $\zeta = \eta(x)e^{(-i\sigma t+iky)}$ and letting $u, v \propto e^{(-i\sigma t+iky)}$

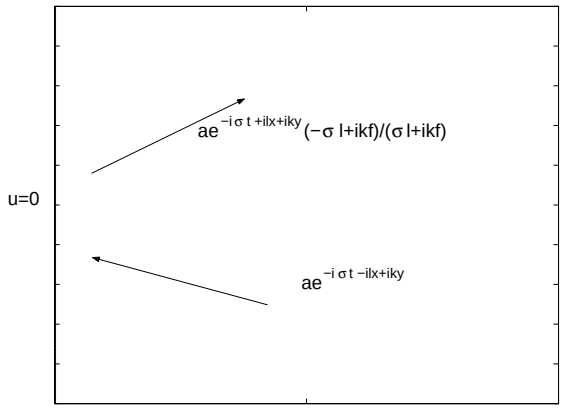


Figure 2: Reflection along a straight coast.

in equations (1),(2),(3) we get

$$i\sigma u - fv = -g\zeta_x, \tag{27}$$

$$i\sigma v + fu = -ikg\zeta, \tag{28}$$

$$i\sigma\eta + (uax)_x + ik(vax) = 0. \tag{29}$$

Eliminating u, v from equations (27),(28) and (29), we get

$$\eta_{xx} + \frac{\eta_x}{x} + \left[\frac{\sigma^2 - f^2}{gax} - \frac{fk}{\sigma x} - k^2\right]\eta = 0. \tag{30}$$

If $\nabla.UD = 0$, then we have

$$\eta_{xx} + \frac{\eta_x}{x} - [\frac{fk}{\sigma x} + k^2]\eta = 0. \tag{31}$$

By making use of the substitution $x = \frac{z}{2k}$ equation (30) can be reduced to the standard Lagurre's Differential Equation

$$\eta_{zz} + \frac{\eta_z}{z} + \left[\frac{\lambda}{z} - \frac{1}{4}\right]\eta = 0, \tag{32}$$

where

$$\lambda = \frac{\sigma^2 - f^2}{2gak} - \frac{f}{2\sigma}. \tag{33}$$

The solution to equation (32) can be written in terms of the eigenfunctions $\eta_n = e^{-z/2}L_n(z), n = 0, 1, 2, ...$, where $L_0(z) = 1, L_1(z) = 1 - 2z, L_2(z) = 1 - 2z + z^2$... are the standard Lagurre's polynomials. The condition that the eigensolutions are finite as $z \to 0, \infty$ gives us

$$\lambda = n + \frac{1}{2}, \qquad\qquad n = 0, 1, 2, \ldots. \tag{34}$$

Substituting for lambda from equation (33) gives us the following dispersion relation,

$$(\sigma^2 - f^2) - \frac{fgak}{\sigma} = (2n+1)gak, \qquad n = 0, 1, 2, \ldots. \tag{35}$$

Notice that if $f = 0$ we have

$$\sigma^2 = (2n+1)gak. \tag{36}$$

If $\nabla.UD = 0$ then

$$\sigma = \frac{f}{2n+1}. \tag{37}$$

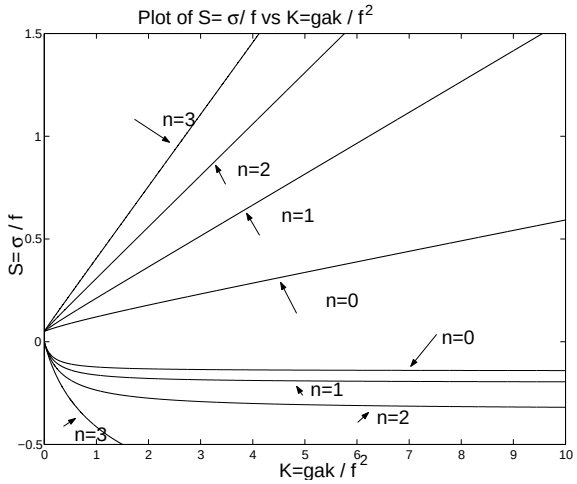


Figure 3: Dispersion relation for waves on a beach with sloping bottom

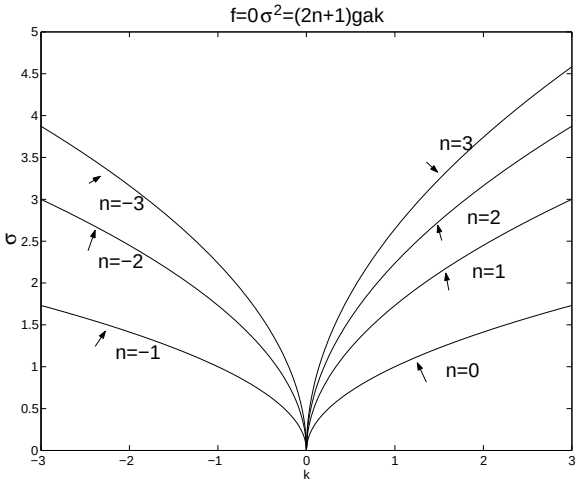


Figure 4: Dispersion relation for waves on a beach with sloping bottom with f=0

7 Refraction of gravity waves over a shelf

A shelf will cause gravity waves to refract. We begin with a shelf of arbitrary topography H(x) (see Figure 5). The local dispersion relation for gravity waves without rotation is

$$\sigma^2 = gH(x)\left[l(x)^2 + k^2\right], \tag{38}$$

where H and l depend on x. Taking the derivative of σ with respect to l and k gives the group velocity in the x and y directions respectively:

$$c_g = (l, k)\sqrt{\frac{gH}{l^2 + k^2}}. \tag{39}$$

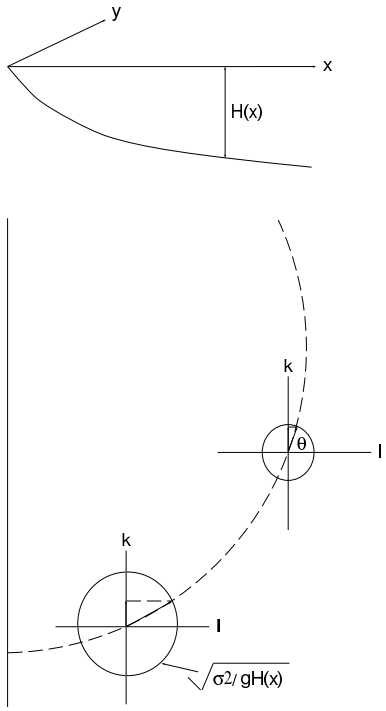


Figure 5: Path of a reflected gravity wave over a sloping shelf.

Equation (39) shows that the group velocity vector is parallel to the wave vector, hence the wave group moves in the direction of the wave vector. As the wave group moves, σ and k don't change, only l does. So if the topography $H(x)$ is known, then l can be computed

in each position and is seen to decrease with distance x for a constant direction slope. Consequently, the direction of the wave vector changes with x. The angle θ that the wave vector makes with the horizontal can be computed using the dispersion relation as

$$\sin\theta = \frac{k}{\sigma}\sqrt{gH(x)}. \tag{40}$$

As H increases, so does θ, leading to the reflection pattern seen in Figure 5. The initial angle with which the wave hits the coastline dictates how far back into the ocean it reflects before returning to the coast again.

The solution for the free surface elevation also shows that all gravity waves are trapped along the coastline with this topography. Using the previous results for short waves with no Coriolis force and an arbitrary depth profile, we get the following expression for η:

$$\eta_{xx} + \left(\frac{\sigma^2}{gH} - k^2\right)\eta = 0. \tag{41}$$

The coefficient of η indicates a turning point in the solution for the free surface. While the coefficient is positive $\left(\frac{\sigma^2}{gH(x)} - k^2 > 0\right)$, a wave solution results. When the coefficient is negative $\left(\frac{\sigma^2}{gH(x)} - k^2 < 0\right)$, an Airy solution results, which decays in the positive x direction.

When the shelf takes the form of a step (see Figure 6), some of the gravity waves will become refractively trapped, but there will be a continuum of waves that will be reflected back into the open ocean. Whether a wave becomes trapped or reflects back is determined by its initial angle of entry onto the shelf.

Munk et al. (1964) investigated refractively trapped gravity waves (or edge waves) along the California coast. Using the actual shelf topography with the assumption of a straight coastline for Southern California, they computed the dispersion relation for the gravest edge wave modes. Their computations matched very well with measurements.[1]

8 Refraction of topographic Rossby waves over a shelf

When a patch of fluid moves on or offshore, potential vorticity must be conserved:

$$\frac{D}{Dt}\left(\frac{\xi + f_0}{H + \eta}\right) = 0. \tag{42}$$

Assuming that $H >> \eta$ and neglecting the nonlinear terms, this implies that

$$\xi_t = u\frac{f_0 H_x}{H}. \tag{43}$$

If there is an initial patch of positive potential vorticity, a positive velocity is induced on one side of the patch, while a negative velocity is induced on the other side. For a shelf sloping in a constant direction ($dH/dx > 0$), Equation (43) states that the vorticity will decrease with the negative velocity and increase with the positive velocity (Figure 7). This propagates the positive vorticity patch in the direction of Kelvin waves.

The dispersion relation for topographic Rossby waves can be derived from analogy with planetary Rossby waves. For planetary Rossby waves, the dispersion relation is

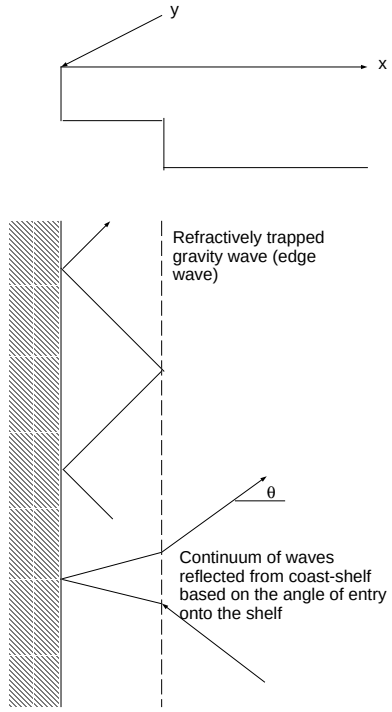


Figure 6: Path of a reflected gravity wave over a step shelf.

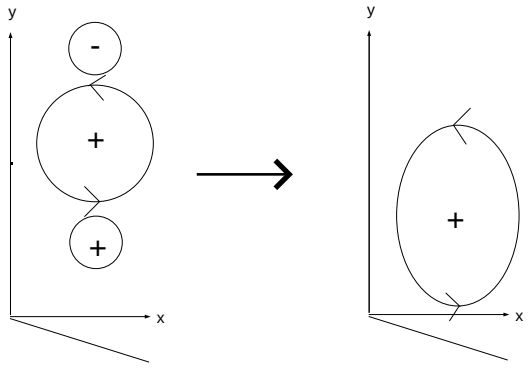


Figure 7: The patches of vorticity generated by an initial patch of positive vorticity.

$$\sigma = -\frac{\beta l}{l^2 + k^2 + f_0^2/gH}. \tag{44}$$

Replacing β with $f_0 H(x)/H$ for topographic Rossby waves gives

$$\sigma = -\frac{f_0 H(x)/H}{l^2 + k^2 + f_0^2/gH}. \tag{45}$$

Similarly to the gravity waves discussed above, the alongshore wavenumber k and the frequency σ do not change as the wave moves in the x direction, so l responds to the change in depth, resulting in the same arcing pattern seen for gravity waves in Figure 5.

9 Wave excitation by tides

Munk et al. (1970) measured tidal pressures and currents off the coast of Southern California and analyzed the data for the tidal components. The dispersion relations for the possible trapped waves against a straight coast with a shelf can be computed. The observations can be fit by superposing all the possible trapped waves of tidal frequencies. Using a plot of the dispersion relations for the various wave types described above (see Figure 8 for a composite illustration of the various dispersion relations), the types of waves that can be excited by a tide is determined by selecting the tidal frequency of interest and seeing which types of waves are possible at that frequency. For example, in Figure 8, if the frequency of interest is f_0, only the Kelvin wave would be excited. If the frequency is very low, the Kelvin and the Rossby waves would be excited. If the frequency is higher than f_0, waves in the Poincaré continuum would be excited in addition to certain edge wave modes and the Kelvin wave. The model of Munk et al. (1970), based on wave excitation, reproduces the observed tidal amplitudes of the M_2 tide, although the K_1 tide is not well modeled. The model was successful in predicting the tidal amplitude very well over a range of model parameters. Additionally the model was also able to predict the amphidrome of the M_2 tide off the coast of Southern California. Comparison of their modeled values for the current over the course of a month shows that the model agrees well with the data, but there are large differences at some points. This is attributed to the inability of the model to resolve baroclinic modes and their currents. [2]

Cartwright (1969) describes measurements which show strong diurnal tidal currents near the St. Kilda islands. The tidal amplitude of the diurnal K_1 tide are much less than the amplitude of the semi-diurnal M_2 tide, but measurements show that the currents due to these tides are of the same order. He attributes this to excitation of a Rossby wave by the diurnal tides.[3]

10 Waveguide modes

In an infinitely long channel in the x direction on an f plane with sides at $y = 0$ and $y = a$ with $v = 0$ at boundaries (Figure 9a), waves will travel and have structure along the waveguide, similar to Kelvin waves propagating along a coast. The kinds of waves which can propagate can be determined by solving

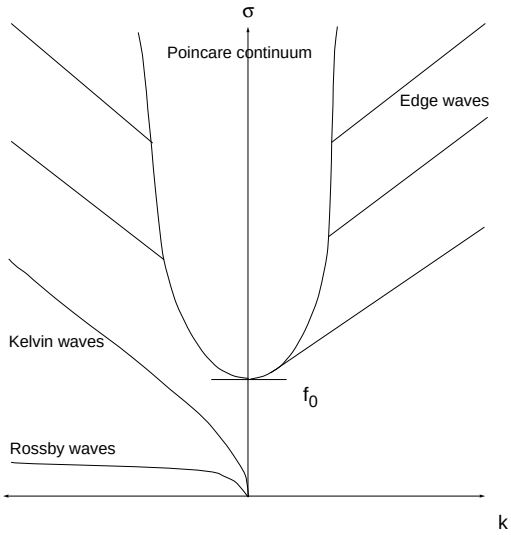


Figure 8: Composite illustration of the dispersion relations for the various wave types discussed above.

$$\nabla^2\eta + \frac{\sigma^2 - f^2}{gD}\eta = 0, \tag{46}$$

where D is the water depth. The boundary condition $v = 0$ can be entered into the governing equations and assuming that v takes a solution form of $e^{i\sigma t+kx+ly}$ the following condition at the boundary results:

$$\sigma\eta_y + f\eta_x = 0 \tag{47}$$

at $y = 0$ and $y = a$. The solution to Equation (46) is assumed to take the form:

$$\eta \sim e^{ikx}\left(\cos\frac{m\pi y}{a} + \alpha_m \sin\frac{m\pi y}{a}\right).$$

This solution form is valid if

$$k^2 = \frac{\sigma^2 - f^2}{gD} - \left(\frac{m\pi}{a}\right)^2. \tag{48}$$

If $k^2 > 0$, the waves will propagate. If $k^2 < 0$ the waves are evanescent and will decay. The boundary condition can be restated as

$$i\sigma\frac{m\pi}{a}\left(-\sin\frac{m\pi y}{a} + \alpha_m\cos\frac{m\pi y}{a}\right) + ifk\left(\cos\frac{m\pi y}{a} + \alpha_m\sin\frac{m\pi y}{a}\right) = 0. \tag{49}$$

at $y = 0, a$. To satisfy the boundary conditions,

$$\sigma_m = -\tfrac{f}{\sigma}\tfrac{ka}{m\pi} \qquad\qquad m = 1, 2, \ldots. \tag{50}$$

Note that the $m = 0$ mode does not satisfy the boundary conditions t $y = a$.

As m increases, k decreases and will become imaginary, making evanescent waves. These waves are unimportant in the infinite channel case, but are important in the closed channel case discussed in the next section. For the waves to propagate

$$m < \left(\frac{\sigma^2 - f^2}{gD} \frac{a^2}{\pi^2} \right)^{1/2} \qquad \text{and} \qquad \sigma^2 > f^2.$$

Kelvin waves are also possible and can propagate along both boundaries, but in opposite directions. The two Kelvin waves must be superimposed to describe the behavior of η. This superposition of the two waves leads to amphidromes in the channel, separated by a half wavelength, where there is no change in η. The wave crests rotate around these amphidromes with a period of $2\pi/\sigma$. Figure 9 shows examples of the location of the amphidromes, depending on the magnitude of the two Kelvin waves.

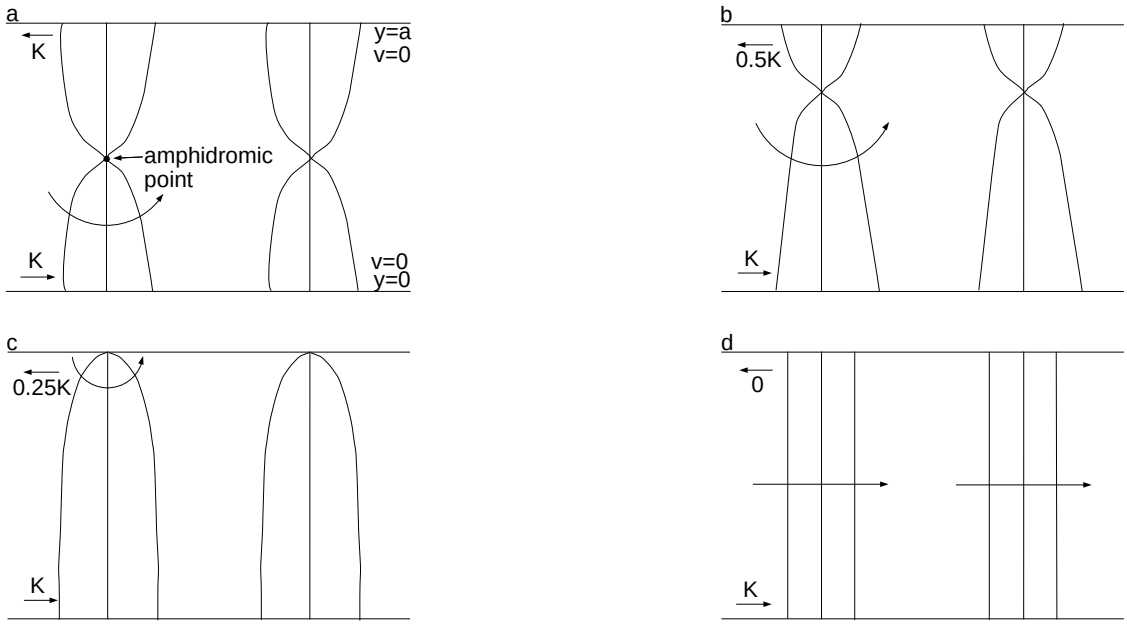


Figure 9: Kelvin waves propagating in a channel. Panel (a) shows the variables. a. The Kelvin waves are of equal magnitude. b. One Kelvin wave is twice the magnitude of the other. c. One Kelvin wave is four times the magnitude of the other. d. The Kelvin wave only exists in one direction resulting in cotidal lines that progress along the channel.

11 Kelvin wave reflection

In a channel with a closed end as in Figure 10, an incident Kelvin wave will be reflected out of the channel. While $v = 0$ at the boundaries for these waves, $u \neq 0$. For the boundary condition of $u = 0$ to be met at the end of the channel, an infinite series of evanescent Poincaré waves (as discussed above) is needed so that the velocity imposed by the Kelvin wave on the end of the channel is countered. These waves decay exponentially

so at a sufficient distance from the end of the channel, the channel is seen as infinite and the solution looks like that discussed above. In order to have these evanescent waves and consequently perfect reflection of the incident Kelvin wave,

$$\left(\sigma^2 - f^2\right) \frac{a^2}{\pi^2 g D} < 1.$$

This can be achieved in two manners: (1) the waves are sub-inertial ($\sigma < f$) or the channel is sufficiently deep or narrow (D or a small, respectively). If neither of these criteria are met, the reflection is not a simple Kelvin wave.

Figure 10 shows the solution for perfect reflection of a Kelvin wave in a flat-bottomed perpendicular wall basin. Modes are obtained for each integral number of Kelvin wavelengths around the circumference of the amphidrome closest to the end of the channel. Bottom topography and shelves make other modes possible as well. Hendershott and Speranza (1971) show examples of this phenomenon. They show that when there is dissipation, the amphidromes move toward $y = a$ in Figure 10. In the Gulf of California, which is shallow near the end of the gulf and hence there is large dissipation, the co-tidal lines show only the incident wave - the reflecting wave has been damped out by the dissipation. In the Adriatic Sea, where the shelf is still deep, the amphidrome is closer to the center of the Sea, implying that the incident and reflected Kelvin waves are of the same amplitude.[4]

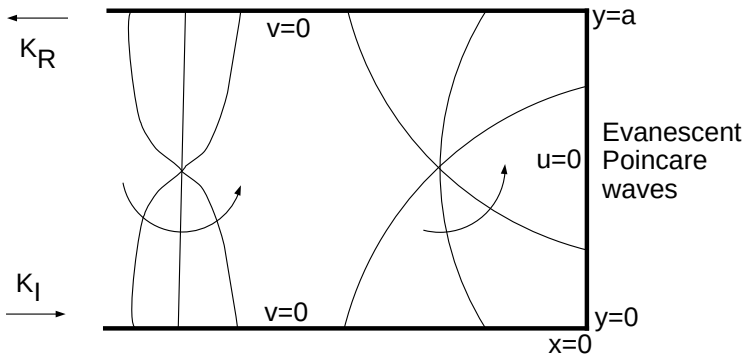


Figure 10: A series of evanescent Poincaré waves is required for reflection of the Kelvin wave out of the closed channel.

12 Tides in gulfs

The conventional treatment of tides in gulfs considers the solution to be the sum of an independent tidal solution (where there is tidal forcing and the tidal amplitude at the mouth of the gulf is zero) and a co-oscillating tidal solution (where there is no forcing and the amplitude at the mouth is observed). Figure 11 shows schematically how this solution is obtained. This method can work only if the tidal amplitude at the mouth of the gulf is observed.

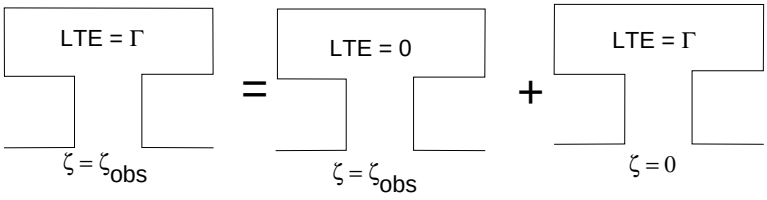


Figure 11: Treatment of tides in a gulf as the sum of an independent and co-oscillating tidal solutions.

Garrett (1975) proposes another methodology to solve this case. Figure 12 shows the two solutions that are combined to achieve the true solution. The boundary condition is that the normal velocity is zero at the coastlines of the gulf and the ocean. Garrett also represents the true solution as the sum of two simpler solutions. In one of these, the zero normal velocity boundary condition is also imposed at the mouth of the gulf. The second solution has a normal mass flux boundary condition at the mouth that is equal to a function $F(s)$ to be determined, where s is the distance across the mouth:

$$\vec{u}D \cdot \hat{n} = F(s), \tag{51}$$

To determine $F(s)$, Garrett first supposes that $\vec{u}D \cdot \hat{n} = \delta(s-\sigma)$ produces elevations $\zeta_G(s) = K_G(s,\sigma)$ and $\zeta_O(s) = -K_O(s,\sigma)$ on the gulf and ocean sides of the mouth respectively, where δ is the Dirac delta function. To satisfy continuity at the mouth, the sum of the two solutions for the gulf and the ocean need to be equal:

$$\zeta(s) = \zeta_G^1(s) + \int_{mouth} K_G(s,\sigma)F(\sigma)d\sigma = \zeta_O^1 - \int_{mouth} K_O(s,\sigma)F(\sigma)d\sigma. \tag{52}$$

To solve this problem, first ζ_G^1,ζ_O^1,K_G,K_O are found and then the total solution,

$$\zeta_G(r) = \zeta_G^1(r) + \int_{mouth} K_G(r,\sigma)F(\sigma)d\sigma, \tag{53}$$

can be solved to give the amplitude of the gulf tide.[5]

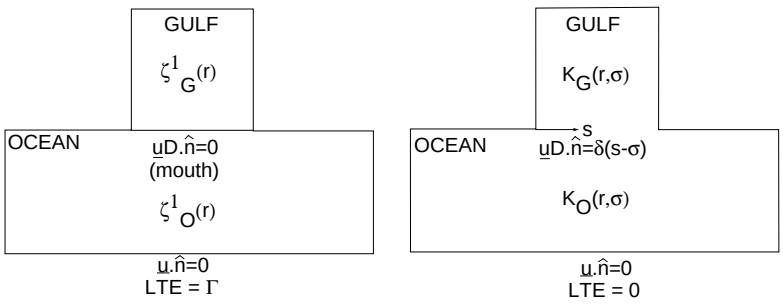


Figure 12: Illustration of the solutions used by Garrett (1975) for the gulf tides.

Notes by Vishwesh and Danielle

References

[1] F. S. W. Munk and F. Gilbert, "Long waves on the continental shelf: an experiment to separate trapped leaky modes," J. Fluid Mech. **20**, 529 (1964).

[2] F. S. W. Munk and M. Wimbush, "Tides off-shore: Transition from california coastal to deep-sea water," Geophysical Fluid Dynamics, year = 1970, volume = 1, pages = 161-235 .

[3] D. Cartwright, "Extraordinary tidal currents near st. kilda," Nature **223**, 928 (1969).

[4] M. Hendershott and A. Speranza, "Co-oscillating tides in long, narrow bays: the taylor problem revisited," Deep-Sea Res. **18**, 959 (1971).

[5] C. Garrett, "Tides in gulfs," Deep-Sea Res. **22**, 23 (1975).

Lecture 6: Internal tides

Myrl Hendershott and Chris Garrett

1 Introduction

So far, we have only considered barotropic tides, which have no vertical structure. In this lecture we turn our attention to baroclinic modes, which do have vertical structure. These are also referred to as *internal tides.*

Internal tides are internal waves forced by the interaction of the barotropic tides with the bottom topography of the oceans. To understand how internal tides can be generated in the ocean, it is useful to go through the properties of internal waves.

2 Internal waves

Internal waves in the ocean are the response of a rotating, density stratified, incompressible fluid to small perturbations. To derive the governing equations for these waves, we begin with the equations of motion for a fluid on a rotating Earth (see lecture 2), use the Boussinesq approximation and linearize about a base state of rest given by $\mathbf{u} = 0$, $\rho = \rho_0(z)$ and $p = p_0(z)$ where $\mathbf{u}$ is the velocity vector, p the pressure and ρ the density. Incompressibility is then

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{1}$$

and the momentum equations become

$$\frac{\partial u}{\partial t} - fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x}, \tag{2}$$

$$\frac{\partial v}{\partial t} + fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y}, \tag{3}$$

$$\frac{\partial w}{\partial t} = -\frac{1}{\rho_0}\frac{\partial p}{\partial z} - \frac{\rho}{\rho_0}g, \tag{4}$$

where f is the Coriolis parameter and ρ and p are the perturbations to the density and pressure fields respectively. The continuity equation becomes

$$\frac{\partial \rho}{\partial t} + w\frac{\partial \rho_0}{\partial z} = 0. \tag{5}$$

And, from the momentum equations, it is also possible to find the energy equation

$$\frac{\partial}{\partial t}\left[\frac{1}{2}\rho_0(u^2 + v^2 + w^2)\right] + \nabla \cdot (p\mathbf{u}) = -w\rho g\hat{k}, \tag{6}$$

where $\hat{k}$ is the unit normal in the z-direction.

From (1)-(5), the equation for internal waves may be derived

$$\frac{\partial^2}{\partial t^2}\left[\nabla^2 w\right] + f^2\frac{\partial^2 w}{\partial z^2} + N^2\nabla_h^2 w = 0, \tag{7}$$

where $N = (-g/\rho_0\,\mathrm{d}\rho_0/\mathrm{d}z)^{-1/2}$ is the Brunt–Väisälä frequency, assumed constant for simplicity in this lecture unless otherwise stated. Subscript h refers to derivatives in the horizontal plane. We refer the reader to [1] for a derivation and we observe as a reminder that we are considering the situation where the hydrostatic approximation is not valid, since gravitational (or buoyancy) force is essential for the motion.

In the process of obtaining (7), the continuity equation (1) is used to substitute the vertical velocity for the horizontal velocities in the horizontal momentum equations (2) and (3). The result is an equation in w with a forcing term dependent on the horizontal variation of pressure

$$\frac{\partial^3 w}{\partial^2 t\,\partial z} + f^2\frac{\partial w}{\partial z} = \frac{\partial}{\partial t}\frac{\nabla_h^2 p}{\rho_0}. \tag{8}$$

If the disturbance is allowed no horizontal dependence, the above equation describes a harmonic oscillation with frequency f.

Another equation for w of second order in time can be obtained by differentiating in time the vertical momentum equation (4) giving

$$\frac{\partial^2 w}{\partial t^2} + N^2 w = -\frac{1}{\rho_0}\frac{\partial^2 p}{\partial t\,\partial z}. \tag{9}$$

Here a harmonic oscillation of frequency N is obtained if the perturbation has no vertical dependence.

The wave equation (7) is obtained by eliminating pressure between these two expressions, and the above remarks suggest that there will be two limiting cases for the motion. This also calls our attention to the fact that (7) has the structure of a double uncoupled harmonic oscillator. We anticipate that the frequencies of these wave motions will depend only on their angle with respect to the x-y plane and to the vertical direction and not on their amplitude. This 'azimuthal' angle determines how the horizontal and vertical forces at play combine to make the restoring force, while the amplitude has no effect on the frequency of a harmonic oscillator.

Equation (7) has an interesting spatial property, which can be better appreciated if we eliminate the temporal dependence by substituting $w = W(x, y, z)e^{-i\omega t}$ into (7). We obtain

$$W_{zz} - \left[\frac{N^2 - \omega^2}{\omega^2 - f^2}\right](W_{xx} + W_{yy}) = 0. \tag{10}$$

If the term in square brackets is positive, then the vertical and horizontal derivatives have different signs, which reflects the anisotropy in space introduced by the vertical stratification. Furthermore the equation is hyperbolic, which allows for energy propagation. If however the term is negative, then (10) is simply Laplace's equation, but for a scale transformation, and does not allow propagation of energy. This is another way of acknowledging the two limiting cases for the motion pointed out earlier: the equation allows for traveling waves with frequencies between f and N.

2.1 Internal wave solutions

We now consider solutions of (10) in an infinite fluid. The equation may be solved in two ways. Firstly, we may solve it assuming a plane wave form of solution and hence we substitute the trial solution

$$W = W_h e^{i(kx+ly+mz)} \tag{11}$$

into (10). This leads to the dispersion relation

$$\omega^2 = \frac{N^2(k^2+l^2)+f^2m^2}{k^2+l^2+m^2}. \tag{12}$$

Substituting the wave solution (11) into the continuity equation we get

$$\mathbf{k}\cdot\mathbf{u} = 0, \tag{13}$$

where $\mathbf{k}$ is the wave vector and $\mathbf{u}$ the velocity vector. That is to say, the wave vector is perpendicular to the displacement of the particles.

To simplify calculations from now on we assume, without loss of generality, that we have orientated the x-axis in the direction of the horizontal component of the wave vector. The dispersion relation (12) then becomes

$$\omega^2 = \frac{N^2k^2+f^2m^2}{k^2+m^2}, \tag{14}$$

and from it we can calculate the group velocity $\mathbf{c}_g = (c_{gx}, c_{gz})$

$$\mathbf{c}_g = \frac{\partial\omega}{\partial\mathbf{k}} = \ [N^2-f^2]\frac{mk}{\omega(k^2+m^2)^2}(m,-k), \tag{15}$$

and we observe that $\mathbf{k}\cdot\mathbf{c}_g = 0$. Hence the group velocity is also perpendicular to the wave vector. Furthermore we can see that the group velocity and wave vector always have opposite vertical components. This means that if phase propagates downwards energy propagates upwards and vice versa.

If we suppose that $\mathbf{x}(t) = (x(t), z(t))$ is the trajectory of the energy carried by the wave, then $\mathrm{d}\mathbf{x}/\mathrm{d}t = \mathbf{c}_g$ and so the trajectory is the curve that satisfies

$$\frac{dz}{dx} = \frac{c_{gz}}{c_{gx}} = -\frac{k}{m}. \tag{16}$$

This means that the energy trajectories are curves along which the independent variables are related by an ordinary differential equation. They are therefore also the characteristics of the wave equation.

Equation (10) may also be solved using characteristics. The solution, in two dimensions, is given by $W(x,z) = F(x+\alpha z) + G(x-\alpha z)$ where F and G are arbitrary functions and α is the ray slope, given by

$$\alpha^2 = \frac{N^2-\omega^2}{\omega^2-f^2}. \tag{17}$$

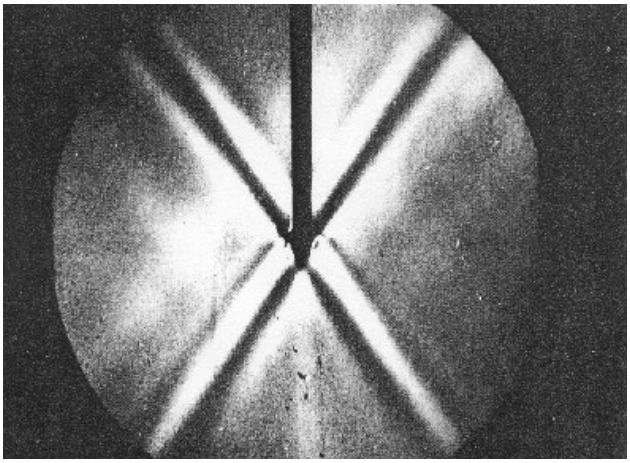


Figure 1: Schlieren image of internal waves, generated in a stratified fluid of constant N, by oscillation of a horizontal cylinder. Note that energy is propagating away along characteristics and surfaces of constant phase stretch out radially from the source. *From Mowbray and Rarity* [2].

If we substitute for ω from the dispersion relation (14) into the expression for α^2, then we find that it reduces to k^2/m^2 and so the result found through this procedure is equivalent to that found in the preceding paragraph.

The unusual properties of internal waves described in this section were observed experimentally by Mowbray and Rarity [2]. They placed an oscillating wave maker in a stratified fluid and observed the wave pattern shown in figure 1.

2.2 Internal waves in a finite depth of fluid

Equation (10) can also be solved for a rigid-lid upper surface and a planar bottom. A rigid-lid is an acceptable approximate boundary condition to use in place of a free surface, as it may be shown that internal waves only have a small surface displacement (see, for example, Pedlosky [1]). The solution is similar to (11), except that the vertical wavenumber is now quantized as

$$W = W_h e^{ikx - i\omega t} \sin\left(\frac{n\pi z}{H}\right), \tag{18}$$

in two dimensions, where H is the depth of the fluid.

This theory predicts the paths of propagation of energy for small perturbations in the ocean, and we see that it depends on the depth of the ocean and on the Brunt–Väisälä frequency. In this analysis, we have assumed that N is constant, however it can be shown (see, for example, Pedlosky [1]) that characteristics are also the ray paths for the WKB approximation for slowly varying N, the case for the ocean.

2.3 Interaction of internal waves with bottom topography

A final interesting property of internal waves is the way in which they interact with a sloping boundary. We refer the reader to Pedlosky [1] for a derivation of the reflection laws. Here we will limit ourselves to stating the results. Frequency is preserved by reflection, but not wavelength. If the slope of the boundary is less than that of the characteristics, the horizontal component of the wave vector is conserved under reflection. In other words, there is no back reflection from a boundary with slope inferior to that of the characteristics. Since all the energy is being transmitted forward, energy density must increase as the waves reach shallower water up the slope. If the slope goes all the way to the surface, energy density would be expected to be infinite at the apex. However, with each reflection the vertical wavenumber gets larger and the waves shorter. Group and phase velocities go rapidly to zero with the depth, so the wave never reaches the apex. In practice, as wavelengths become too small, the linearity assumption breaks and dissipative process come into play. Note also that the closer the slope of the surface is to that of the characteristics, the more intense the focusing effect, as can be seen in figure 2a.

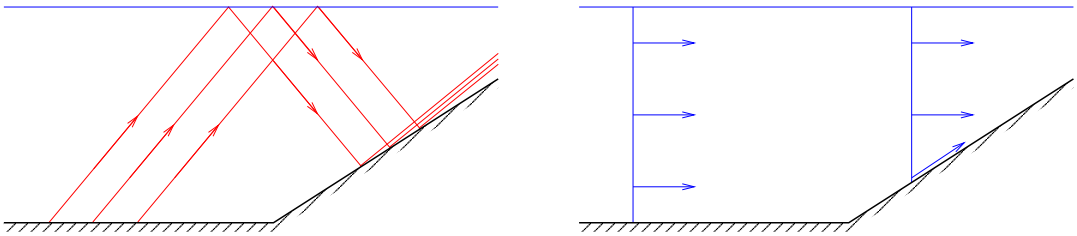


Figure 2: (a) Internal wave characteristics approaching a sloping boundary. After reflection, the energy in the waves is concentrated into a narrower band. If the boundary has slope equal to that of the characteristics, then all characteristics after reflection lie on the same line. (b) A barotropic velocity profile impinging on a sloping boundary. On the boundary, the velocity vector can only have a component tangential to it and hence over the slope, the velocity profile is no longer barotropic.

This behaviour of internal waves in interaction with a sloping boundary was observed in Sandstrom's [3] experiments: the amplification from a sloping bottom was clearly demonstrated.

3 Internal waves in the oceans

It is nowadays accepted that the interaction of the barotropic tide with bottom topography in the oceans is one source of internal waves, in this case the waves are referred to as internal tides. One way to conceptualize this generation is to imagine that, instead of a tidally oscillating ocean with a fixed bottom, the water of the ocean is fixed and the bottom oscillates with tidal frequency. The oscillations of topography then behave as oscillating wave makers, as in the experiment of Mowbray and Rarity described in §§2.1.

However, in the past, the curious behaviour found for the solutions of (7) led to doubts about whether it was in fact a suitable equation to describe the response of the oceans to small perturbations. Another question of interest was whether these waves could be generated in the ocean by the barotropic tides. It was imagined that a sloped boundary could act as a source of internal waves, in the case of a barotropic impinging disturbance. As illustrated in figure 2b, a barotropic disturbance cannot meet the zero normal component requirement at the slope and remain barotropic. Its energy must go into generating internal waves, and propagate away along characteristics. These two questions motivated a number of numerical and observational efforts, some of which we will briefly be described below.

Figure 3 shows the result of a numerical calculation where a barotropic tide was made to impinge on a step-like slope. No specific mode structure can be recognized in the reflected perturbation, because a composition of many is necessary to conform to such an abrupt boundary. However, the results clearly show that the reflected perturbation is baroclinic and propagates along the characteristics.

Regal and Wunsch [5] attempted an observational verification of propagation of energy along characteristics for a real slope, shown in figure 4a, with the measured profile of N shown in figure 4b. They found that the shape of the boundary coincided with that of the characteristics. This meant that intense focusing of energy would occur along the slope. After bottom reflection, this energy would travel to point D on the surface. An energy profile below point D does in fact show concentration near the surface.

Even once it was reasonably established that internal waves can be generated by the interaction of the barotropic tides with topographic features, it remained to be clarified whether these motions had only local influence over ocean dynamics or if they could propagate away for large distances and be ubiquitous in the ocean. This question was addressed using data of an experiment originally designed to study mesoscale circulation in the ocean.

The Mid-Ocean Dynamics Experiment (MODE) provided data that proved very useful for the study of internal waves. The setup consisted of a two dimensional array of current, temperature and pressure profiling moorings as shown in figure 5a. These were arranged in three concentric circles with radii 50, 100 and 180 km, situated over a smooth abyssal plane of depth 5400 km. Being 700 km away from the nearest large topographic feature, this array was far from any source of internal tides. If significant signals of internal waves were detected, then they would have traveled far to get there.

The two dimensional arrangement of moorings could be used as an antenna, the difference in arrival times of a given phase surface in different moorings could be used to infer the direction of propagation of the signal. Internal waves with M_2 and S_2 frequencies were detected coming from the general direction of the nearest major topographic feature, 700 km away. The period and direction of incidence of these signals were strong indications that they were in fact internal tides generated at the Blake Escarpment. Averages over the whole data collecting period for each mooring showed very significant baroclinic contribution to their structure in some cases, indicating that the internal waves were an important part of the local dynamics, as shown in figure 5b.

These results showed that internal tides, once generated, could travel large distances and have a significant influence over the dynamics of the open ocean. They also suggested that a significant portion of the barotropic tidal energy was going into baroclinic tides, since baroclinic and barotropic contributions seemed to have comparable amplitudes in some of

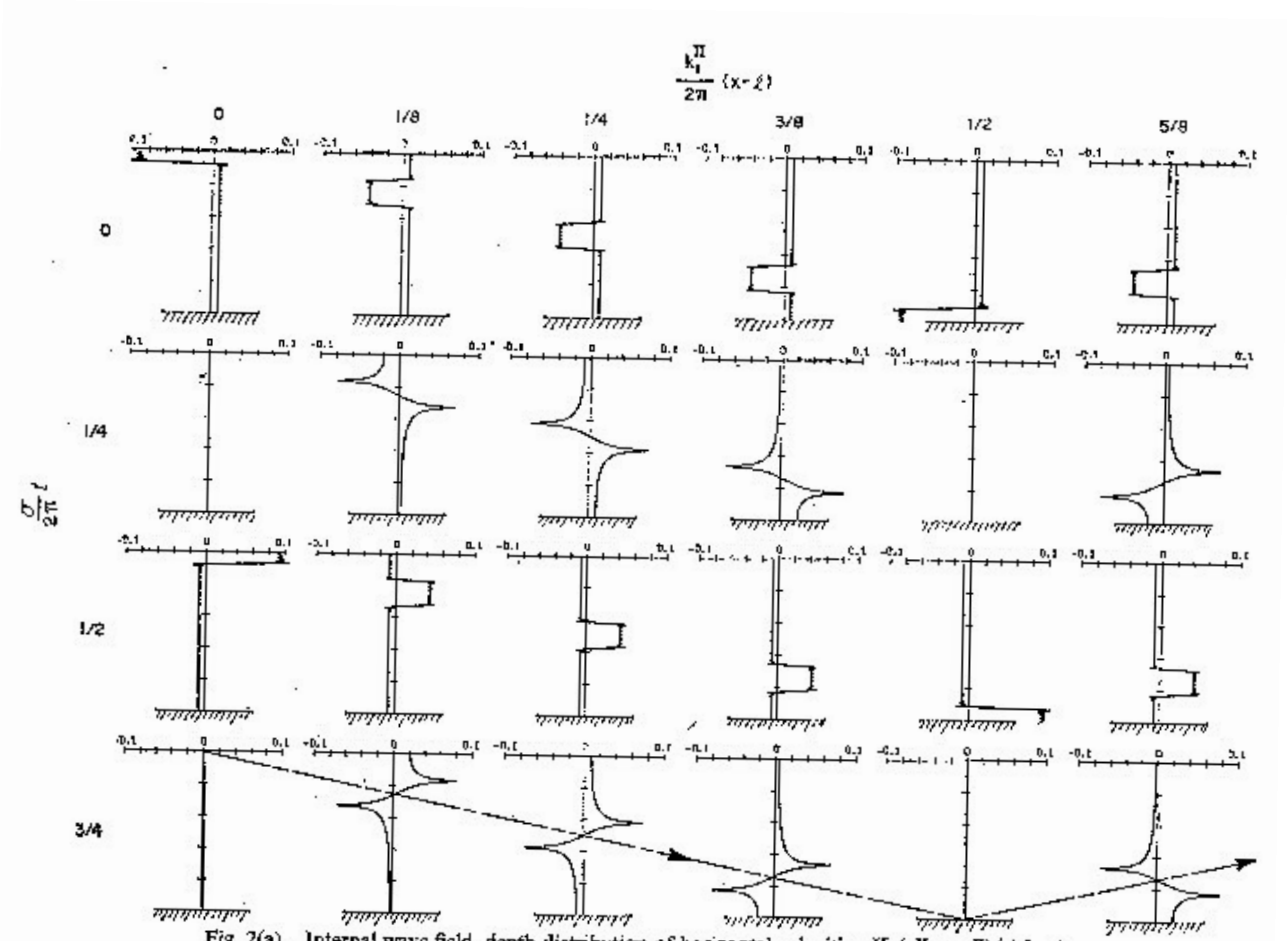


Fig. 2(a). Internal wave field, depth distribution of horizontal velocities, $H_1\,(u^{II}-u_0^{II})/A_0^I$, at position $k_1^{II}(x-l)/2\pi$ and time $\sigma t/2\pi$. Model parameters: $(N^2-\sigma^2)/(\sigma^2-f^2)=715$; $H_1/H_2=1/12{\cdot}5$; $k_0^I H_1=4{\cdot}22\times10^{-4}$; 100 normal modes; antiresonant condition: $k_1^{II}l/2\pi=4{\cdot}50$.

Figure 3: Depth distribution of horizontal internal tidal currents at increasing distances away from a step shelf (top) at two times (centre and bottom) separated by a quarter-wave period. The profile is seen to follow the rays (arrows in bottom) away from the slope. *From Rattray et al.* [4].

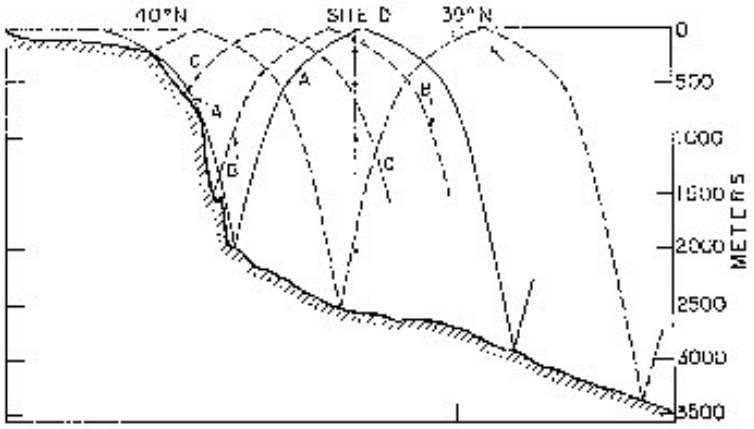


Fig. 4. The topography of the continental margin to the north of Site D, indicating the two families of characteristics originating along the continental slope and passing through the mooring. Notice that one characteristic, marked 'A', is virtually tangent to the slope and eventually reaches the surface near Site D. Characteristics shown as dashed, two of which are marked 'B' and 'C', also tend to focus near the surface (adapted from FOFONOFF, 1966).

(a)

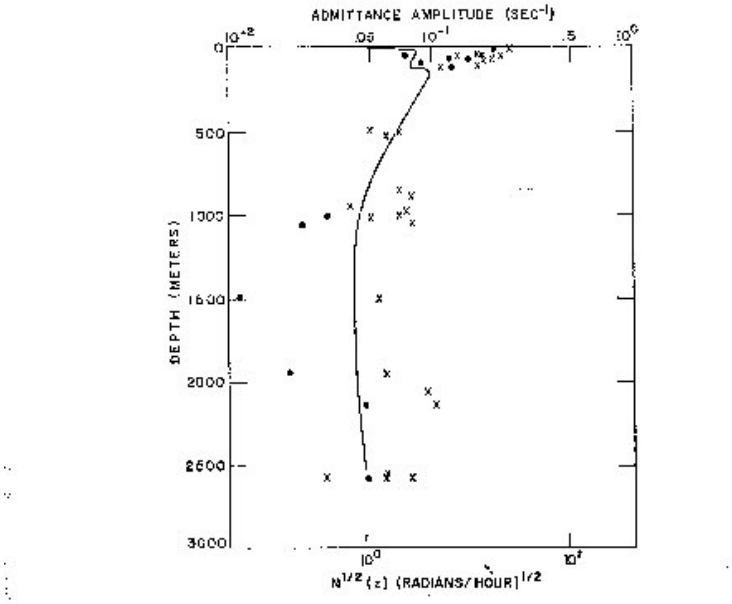


Fig. 2. Admittance amplitudes U and V as a function of depth at Site D. The curve indicates the average square root buoyancy frequency at this location. Though the position of this $N^{1/2}(z)$ curve is arbitrary, we have chosen it to pass through the bulk of the points, clearly showing the near surface intensification. (Cross is U, dot is V.)

(b)

Figure 4: (a) Profile of relief along 70° W together with selected semi-diurnal characteristics passing near site D. (b) Admittance amplitude × for semi-diurnal tidal currents together with Brunt–Väisälä frequency $N(z)$, solid line, at site D. Near surface admittances are strongly intensified. *From Regal and Wunsch* [5].

the mooring averages.

Probably the main reason why the role of baroclinic tides in ocean dynamics was only recently fully acknowledged is their small effect on surface elevation. Typically of only a few centimetres, surface elevation due to internal tides was historically dismissed as noise in satellite data. However the shortcomings in the attempts to infer underwater currents from surface displacement, illustrated in figure 6 are evidence of their role.

Surprisingly it is the small effect on surface elevation that has recently made global coverage possible in internal tide observation, because these small displacements have been successfully observed by satellites, as shown in figure 7.

4 Mixing and internal tides

Internal tides are currently receiving renewed attention due to their role in ocean mixing. In the absence of mixing, the temperature and salinity profiles of the ocean would be uniform, barring a thin surface layer heated by the Sun. Cold, dense water generated at the poles would sink and flow equatorward and upwell. Eventually, water of this property would fill the oceans up to the depth at which solar radiation penetrates. However, the real ocean does not have this structure, as shown in figure 8: vertical mixing prevents its realization.

Internal tides are one possible source of mixing. Two studies whose results support this statement were carried out in the Brazil Basin and the Monterey Canyon off the coast of California. Figure 9a shows levels of turbulent diffusivity across the Brazil Basin. Enhanced diffusivity, and hence mixing, is seen over the rough topography compared to over the smooth topography. This suggests the involvement of bottom topography in ocean mixing, at least at abyssal depths. Both tidal flow, and mean or eddy flows, over topography generate internal waves. However, in this data, a modulation of the dissipation rate over the spring-neap cycle suggests the waves are internal tides.

Figure 9b shows levels of turbulent dissipation along a section extending off the slope of the continental shelf. Enhanced levels are seen along the ray path of a semi-diurnal tide beam, out to more than 4 km away from the topography. Where internal waves dissipate and cause mixing is not really known. It is thought that non-linear wave-wave interactions and scattering when internal waves reflect off the ocean floor cause low vertical wavenumber modes to cascade to higher wavenumbers. At higher vertical wavenumbers, there is increased vertical shear and eventually a shear instability, and hence mixing, results.

However, low wavenumber modes are relatively stable to wave-wave interactions. In scattering also, much of the low modes are preserved and so low modes tend to survive many bottom encounters. This results in a general persistence of low mode internal tides, and they have been observed to propagate to $O(1000\,\mathrm{km})$ from their source, as described in §3.

5 Conversion of energy from barotropic to internal tides

As a first step towards understanding the contribution of internal tides to ocean mixing, research has been conducted into finding the amount of energy converted from barotropic

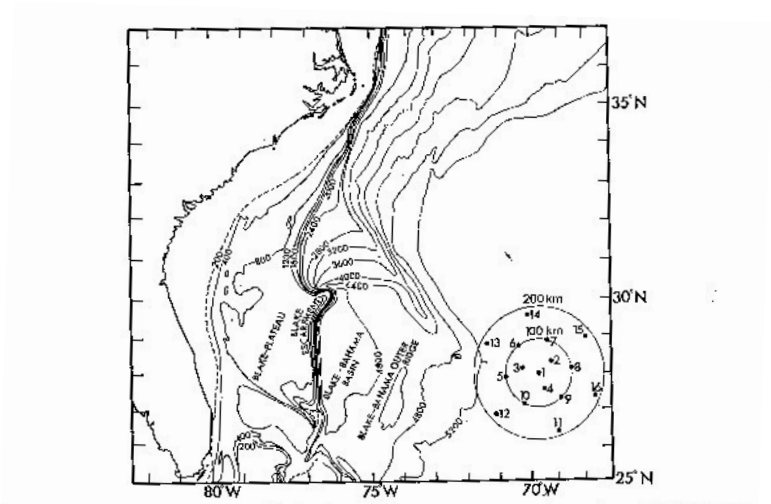


FIGURE 1. Chart of the MODE area, showing the array of fixed moorings. Bathymetric contours in metres.

(a)

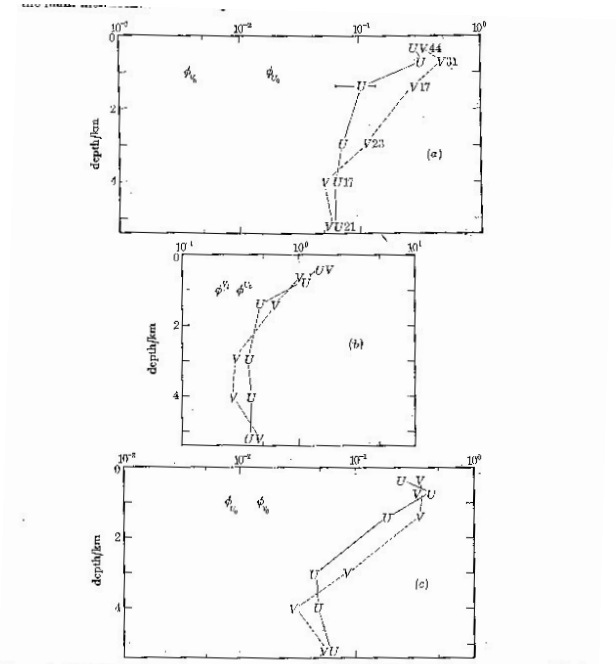


FIGURE 3a. Vertical profile of squared horizontal current ((cm/s)²) for U (east) and V (north) in the S_2 band, averaged at depth levels over the entire array. Estimates of squared amplitude for the barotropic current components U_0 and V_0 are given, showing that the currents are dominated by internal waves at all depths.

FIGURE 3b. Similar estimates for the M_2 band currents. Here the deep currents are greatly influenced by the barotropic mode.

FIGURE 3c. Similar estimates for the N_2 band. Internal waves appear to dominate at all depths.

(b)

Figure 5: (a) Chart of the MODE area, showing the array of fixed moorings. Bathymetric contours are in metres. (b) Vertical profile of the squared horizontal current ($\mathrm{cm^2 s^{-2}}$) for U (east) and V (north) in the (a) S_2 and (b) M_2 band, average at depth levels over the entire array. Estimates for the squared amplitude for the barotropic current components U_0 and V_0 are also given. For the S_2 band, the current is seen to be dominated by internal waves at all depths. For the M_2 band, the deep currents are greatly influenced by the barotropic mode. *From Hendry* [6].

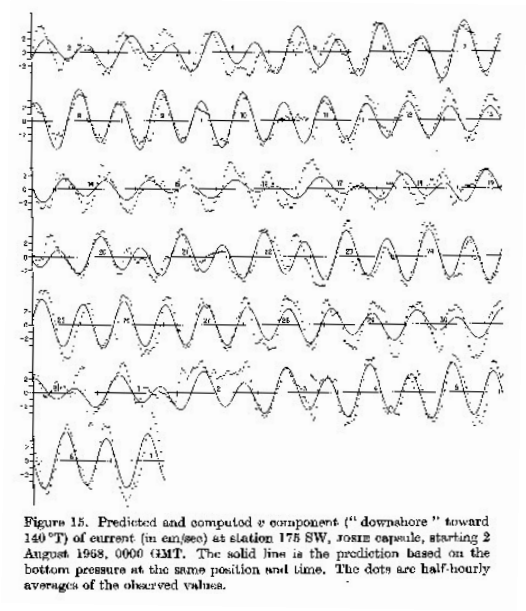


Figure 15. Predicted and computed v component ("downshore" toward 140 °T) of current (in cm/sec) at station 175 SW, JOSIE capsule, starting 2 August 1968, 0000 GMT. The solid line is the prediction based on the bottom pressure at the same position and time. The dots are half-hourly averages of the observed values.

Figure 6: Observed (dotted) and predicted barotropic (solid) longshore bottom velocity off the southern California coast. *From Munk et al.* [7]

.

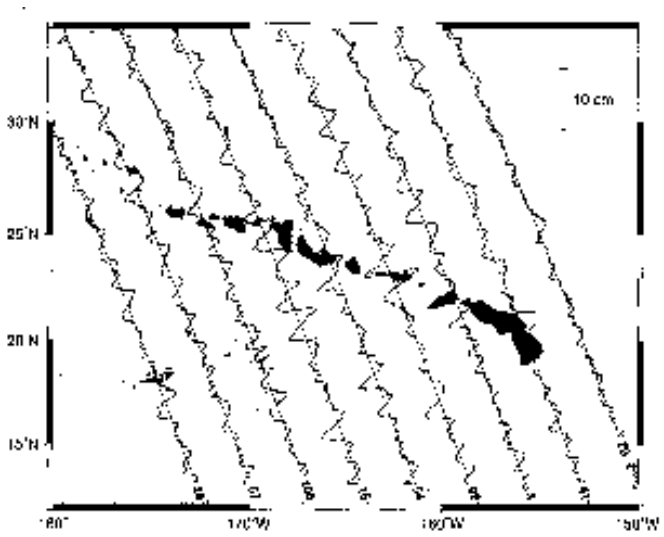
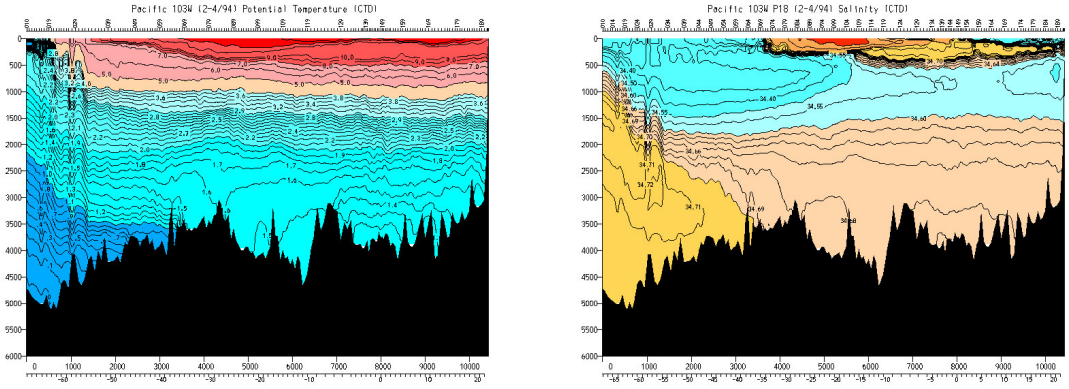


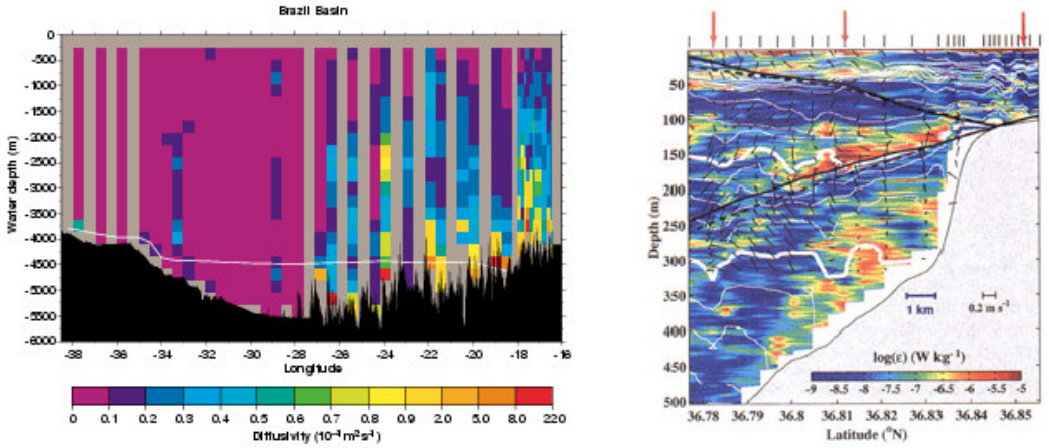
Figure 7: High-pass filtered estimates of M_2 tidal amplitudes plotted along Topex–Poseidon ground tracks. Background shading corresponds to bathymetry, with darker denoting shallower water. *From Ray and Mitchum* [8]

.



(a) (b)

Figure 8: (a) Potential temperature and (b) salinity profiles along a section of the Pacific Ocean from approximately 70° S to 20° N. Red indicates high values and blue low. The high salinity values at the surface are due to evaporation. *From the World Ocean Circulation Experiment.*



(a) (b)

Figure 9: (a) A section of turbulent diffusivity across the Brazil Basin. The bottom bathymetry is shown from a representative transect with the apex of the Mid-Atlantic Ridge at 12° W. *From Toole et al.* [9]. (b) A section of turbulent dissipation rate ϵ off the shelf break slope in Monterey Bay, California. The solid black lines are the ray paths of semi-diurnal tide beams. *From Lien and Gregg* [10].

to baroclinic tides. As discussed in §3, the interaction of the barotropic tide with bottom topography in the oceans results in the generation of internal tides.

Cox and Sandstrom [11] and Bell [12] developed linear[1] theories to model the generation of internal tides from the barotropic tide. Rough estimates originally suggested that the energy flux from barotropic to internal tides (contributing to energy loss in the deep oceans) was much less than the energy dissipated in shallow seas, as shown in figure 10. However, even if only 10% of the total energy loss of the barotropic tides took place in the deep oceans, it would contribute significantly to abyssal mixing.

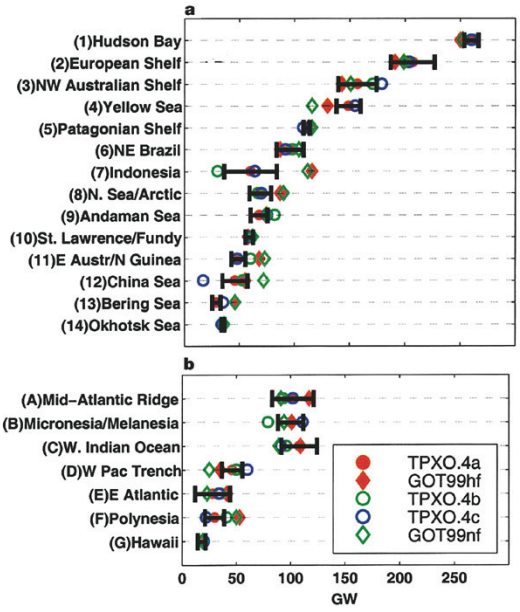


Figure 10: Area-integrated dissipation for various shallow seas and deep-ocean areas. *From Egbert and Ray* [13]

.

Egbert and Ray [13, 14] examined least-squares fits of Topex–Poseidon altimetry data to models for the global barotropic tide. They interpreted model residuals in terms of 'tidal dissipation' with this referring to any mechanism that transfers energy out of the barotropic tides: they were not able to distinguish whether barotropic energy is lost to baroclinic modes or to bottom friction. Figure 11 shows the regions of significant 'tidal dissipation'.

About 0.7 TW of power is lost from M_2 barotropic tides in the deep ocean. The fraction of this associated with frictional dissipation can be estimated from the drag force $\rho c_D |u| u$ where $c_D = 0.0025$ is the drag coefficient. For typical open ocean tidal speeds of $u \simeq 0.03\,\mathrm{ms}^{-1}$, the frictional dissipation is less than $0.1\,\mathrm{mWm}^{-2}$, or less than 30 GW globally. Thus nearly all barotropic tidal loss in the deep ocean must occur as internal tide generation.

One possible location for significant internal tide generation is at continental slopes. In some places, such as the north-western Australian shelf and the Bay of Biscay, generation

[1]Linear will be defined more precisely in §§5.1. Essentially, it assumes that the heights and slopes of topographic features are small.

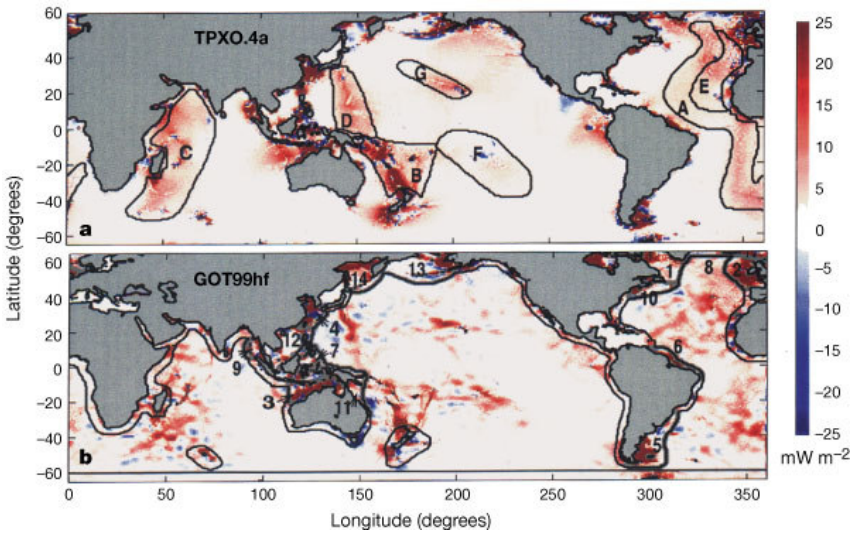


Figure 11: Global distribution of barotropic energy loss. The estimates were made using observations of sea surface elevation from the Topex–Poseidon altimeter. Positive values (red) show energy loss. Negative values (blue) indicate regions where noise prevented accurate estimates. *From Egbert and Ray* [14].

at the continental shelf is significant because the tidal flow is perpendicular to the shelf, as in figure 12a. However, in many places tidal flow is parallel to the coast, as in figure 12b and hence internal tide generation is small. For a few decades, this area was the main focus of research into internal tides.

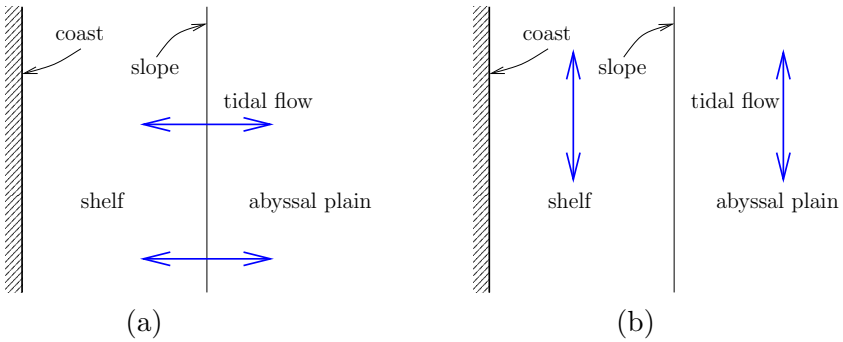


Figure 12: The approximate flow of the barotropic tide close across the continental shelf. (a) Flow perpendicular to the slope and (b) flow parallel to the slope.

The other main possibility for significant internal tide generation is at mid-ocean ridges and ocean island chains such as the Hawaiian Ridge. Recently, research has focused on this aspect of internal tide generation. Following this work, we calculate the energy conversion between the barotropic tide and the internal tides for different topographies. For this, we consider two approaches: linear theory for small slope, small height topography and a full

theory for infinitely steep topography. In both approaches, the Brunt-Väisälä frequency is assumed constant and we restrict ourselves to two-dimensional problems.

5.1 Linear theory for energy conversion

Consider a barotropic tide $U_0 \cos \omega t$, where ω is an astronomically determined forcing frequency (for example the frequency of the M_2 or S_2 tide), over rough topography as shown in figure 13.

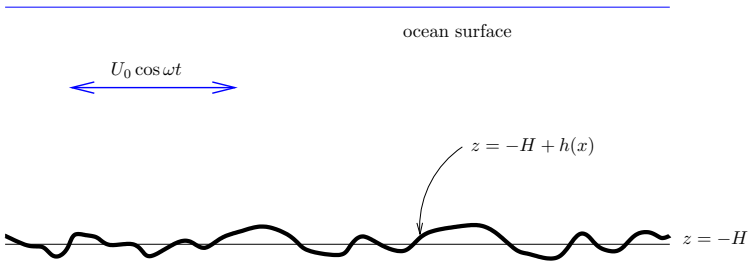


Figure 13: A barotropic tide $U_0 \cos \omega t$ across a rough bottom $z = -H + h(x)$ where $h(x)$ is a bottom topography relative to a deep constant depth level $z = -H$.

We first look at the three key dimensionless parameters of the problem.

- $\frac{\text{tidal excursion}}{\text{horizontal scale of topography}} = \frac{U_0 k}{\omega}$ where k is a wavenumber associated with the horizontal scale of the topography;
- $\epsilon = \frac{\text{maximum bottom slope}}{\text{slope of characteristics}}$ with $\epsilon < 1$ referred to as *subcritical* as the rays have steeper slope than the bottom topography, $\epsilon = 1$ is *critical* and $\epsilon > 1$ is *supercritical*;
- $\delta = \frac{\text{maximum topographic height}}{\text{ocean depth}} = \frac{h_0}{H}$ where h_0 is an amplitude associated with the bottom topography and H is the depth of the ocean.

In addition there are two dimensionless numbers ω/f and ω/N which specify the time-scale of the tide relative to the inertial and buoyancy time scales.

5.1.1 Assumptions

Firstly, it is assumed that the tidal excursion is much smaller than the horizontal scale of the bottom topography. If this were not the case, then the topography would essentially see a nearly constant current across it, although switching direction over the tidal cycle. Quasi-steady lee waves would then form on the downstream side of the topography. For M_2, and with deep ocean currents of the order $10^{-2}\,\mathrm{ms}^{-1}$, the tidal excursion is of the order of 100 m and so $U_0 k/\omega$ is indeed small.

Secondly, the linear theory requires that ϵ and δ are also small. Note that as the wavenumber k associated with the bottom topography gets higher, ϵ increases since the characteristic slope, α from (17) is fixed by fixing ω but the slope of the topography is proportional to k. Hence, for sufficiently high wavenumbers, the linear theory will always get into trouble. From observations by St. Laurent and Garrett [15], over 90% of the energy flux is accounted for in wavenumbers below critical at mid-ocean ridges.

5.1.2 Energy flux

At the bottom, the boundary condition of zero normal flow, in linearized form, becomes

$$w|_{z=-H} = U_0 \frac{\mathrm{d}h}{\mathrm{d}x},$$

where w is the vertical component of the baroclinic wave response, and so $w \propto k$. From the dispersion relation for internal waves (14), it can be seen that the vertical component of the group velocity for a rigid-lid solution (18) is

$$c_{gz} = \frac{\partial \omega}{\partial (n\pi/H)} = -\sqrt{\frac{(\omega^2 - f^2)^3 (N^2 - \omega^2)}{\omega^2 (N^2 - f^2)^2}} k^{-1},$$

and so $c_{gz} \propto k^{-1}$ since ω is a fixed forcing frequency.

Hence the kinetic energy is proportional to k^2 and the vertical energy flux to internal tides has the form

$$F_{\text{linear}} \propto c_{gz} \times w^2 \propto k \times \text{bottom topography spectrum},$$

where the subscript 'linear' implies we are considering the linear regime. We can use Fourier decomposition and superposition to obtain solutions for arbitrary topography.

Llewellyn Smith and Young [16] showed that the horizontal energy flux away from the topography is in fact

$$F_{\text{linear}} = F_0 \alpha \pi H^{-3} \sum_{n=1}^{\infty} k_n \tilde{h}(k_n) \tilde{h}^*(k_n), \tag{19}$$

where the Fourier transform is defined as

$$\tilde{h}(k_n) = \int_{-\infty}^{\infty} h(x) e^{-ikx} \mathrm{d}x,$$

and the modes are given by $k_n = \alpha \pi n / H$, from (14), where α is the ray slope given in (17) and

$$F_0 = \frac{1}{2\pi} \rho_0 \omega^{-1} \sqrt{(N^2 - \omega^2)(\omega^2 - f^2)} U_0^2 H^2.$$

We now consider two different forms for the topography and calculate the associated energy fluxes.

5.1.3 A small, subcritical ridge

Assume a bottom topography given by a Witch of Agnesi curve[2]

$$h(x) = h_0 \left(1 + \frac{x^2}{b^2}\right)^{-1},$$

as shown in figure 14a. Here b parameterizes the width of the topography.

[2]Originally this curve was given the Latin name *versoria* which means 'rope that turns a sail' from the construction resulting in this curve. This became the Italian *la versiera*, but on translation into English, of a textbook by Maria Agnesi, it was mistaken for *l'aversiera*, or 'witch'.

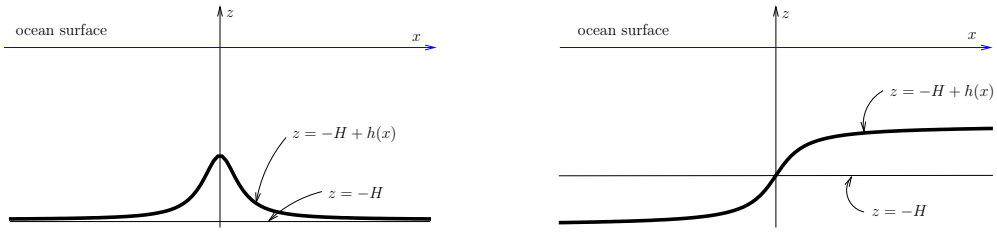


Figure 14: The profiles for (a) a small, subcritical ridge and (b) a small, subcritical step.

Provided that $\delta \ll 1$ and the maximum slope of the Witch, $(3^{3/2}/8)(h_0/b)$, is sufficiently small so that $\epsilon \ll 1$, the problem is linear. The Fourier transform of the Witch is $\tilde{h}(k) = bh_0\pi e^{-kb}$ and hence

$$F_{\text{linear}} = F_0 \frac{\pi^2}{4}\delta^2 \sum_{n=1}^{\infty} nc^2 e^{-nc},$$

where $c = (3^{3/2}\pi/4)(\delta/\epsilon)$ and $\epsilon = \alpha^{-1}(3^{3/2}/8)(h_0/b)$. Hence we have an expression for the energy loss to internal tides.

One further question which may be asked is how strongly the result depends on the width of the topographic feature. In the small δ/ϵ limit $F_{\text{linear}} \sim F_0\pi^2\delta^2/4$, hence in the deep ocean the flux is independent of the ridge width.

This type of profile is appropriate for mid-ocean ridges which have subcritical slopes and are relatively small compared with the overall depth of the ocean.

5.1.4 A small, subcritical step

The bottom topography is given by

$$h(x) = \pi^{-1}h_0 \tan^{-1}\frac{x}{b},$$

as shown in figure 14b. It should be noted that the derivative is proportional to the Witch profile and, using this fact and integration by parts, the Fourier transform becomes $\tilde{h}(k) = -ih_0k^{-1}e^{-kb}$. The flux is now

$$F_{\text{linear}} = F_0\delta^2 \sum_{n=1}^{\infty} n^{-1}e^{-2n\delta/\epsilon},$$

where $\epsilon = h_0/(b\pi\alpha)$.

For small δ/ϵ, we obtain

$$F_{\text{linear}} \sim F_0\delta^2 \log(\epsilon/2\delta). \tag{20}$$

Hence, unlike the ridge in §§5.1.3, the flux remains dependent on the width of the sloping region, even in a deep ocean.

It should be noted, that it is not possible to approximate arbitrary topography by a series of independent steps, as was done by Sjöberg and Stigebrandt [17]. The flux at each

step, if isolated and independent, is proportional to $\delta^2 \log \delta$ as this will be shown later to apply even for an abrupt step. If there are n steps, then $\delta \propto h_0/n$ and the total energy flux is $F_{\text{linear}} \propto n \times h_0^2 n^{-2} \log n$, which tends to 0 as n tends to infinity.

5.2 Energy conversion at supercritical topography

For supercritical topography, $\epsilon > 1$ and the linear theory of the previous section is no longer valid. In this case, it is possible to use a Fourier series decomposition and matching to study some limiting cases: a knife-edge ridge and a Heaviside step.

5.2.1 A knife-edge ridge

The topography in this case is assumed to consist of a knife-edge of height h_0, as shown in figure 15a.

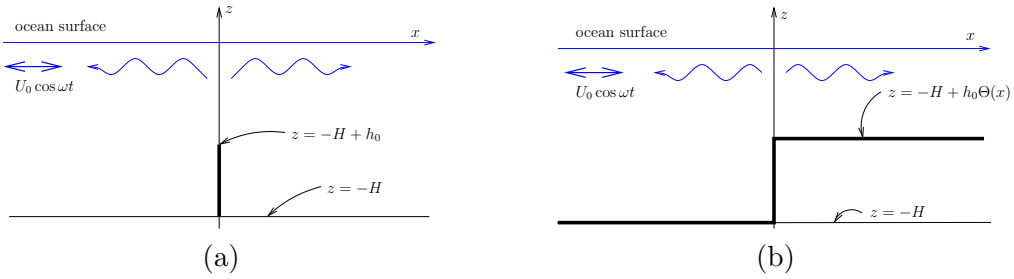


Figure 15: The profiles for (a) a knife-edge ridge and (b) a Heaviside step.

Note that the ratio of the tidal excursion to the horizontal scale of the topography will now necessarily be infinite, and quasi-steady lee waves should be expected. However we ignore these as the same analysis with a top-hat ridge gives very similar results (see St. Laurent *et al.* [18]).

The barotropic tidal current is given by $U_0 \cos \omega t$, perpendicular to the ridge. When incident on the knife-edge, it produces waves propagating to the left and to the right, at the same frequency as the barotropic tide. The form for the velocities in the reflected baroclinic modes are taken as

$$u_1 = U_0 \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi z}{H}\right) \cos(k_n x + \omega t), \quad w_1 = \alpha U_0 \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi z}{H}\right) \sin(k_n x + \omega t) \quad (21)$$

to the left of the knife and

$$u_2 = U_0 \sum_{n=1}^{\infty} b_n \cos\left(\frac{n\pi z}{H}\right) \cos(-k_n x + \omega t), \quad w_2 = -\alpha U_0 \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi z}{H}\right) \sin(-k_n x + \omega t) \quad (22)$$

to the right, where α is the characteristic slope from (17). The form for w_i is assumed from (18) and the form for u_i follows from continuity equation (1). The form for the pressures may be derived using (21) and (22) and the linearized momentum equations (2) and (4) with $f = 0$.

There are matching conditions at $x = 0$ given by

$$w_1 = w_2, \qquad -H + h_0 \le z \le 0, \tag{23}$$

$$u_1 + U_0 = u_2 + U_0, \qquad -H + h_0 \le z \le 0, \tag{24}$$

$$u_1 + U_0 = 0, \qquad -H \le z \le -H + h_0, \tag{25}$$

$$u_2 + U_0 = 0, \qquad -H \le z \le -H + h_0. \tag{26}$$

The latter three conditions combine to give

$$\sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi z}{H}\right) = \sum_{n=1}^{\infty} b_n \cos\left(\frac{n\pi z}{H}\right), \qquad -H \le z \le 0,$$

which implies, from orthogonality of cosines, that $a_n = b_n$.

The remaining conditions then become

$$\sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi z}{H}\right) = 0, \qquad -H + h_0 \le z \le 0,$$

from (23) and

$$\sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi z}{H}\right) = -1, \qquad -H \le z \le -H + h_0,$$

from (25) or (26).

Multiplying by $\cos(m\pi z/H)$ and integrating vertically gives

$$\sum_{n=1}^{\infty} \left(\int_{-H}^{-H+h_0} \cos\left(\frac{n\pi z}{H}\right) \cos\left(\frac{m\pi z}{H}\right) \mathrm{d}z + \int_{-H+h_0}^{0} \sin\left(\frac{n\pi z}{H}\right) \cos\left(\frac{m\pi z}{H}\right) \right) \mathrm{d}z = - \int_{-H+h_0}^{0} \cos\left(\frac{m\pi z}{H}\right) \mathrm{d}z.$$

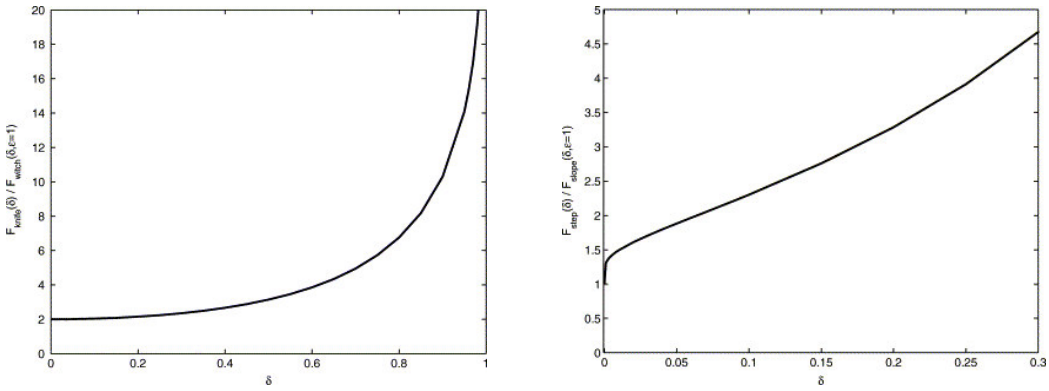
Curtailing the summation at M and taking $m = 0, 1, 2, \ldots, M-1$ gives a matrix equation to solve for a_n of the form $\mathsf{A}_{mn} a_n = c_m$ which may be solved numerically.

The flux of energy into the baroclinic modes may now be calculated. The conversion rate is equal to the energy flux away from the topography. From (6), this may be written as

$$F_{\text{knife}} = \int_{-H}^{0} < pu > \, \mathrm{d}z,$$

where $< \cdot >$ is the average in time. Hence $F_{\text{knife}} = F_0 \sum_{n=1}^{\infty} n a_n^2$.This flux is appropriate for calculating the flux from ocean island chains, such as Hawaii, since slopes at such features are supercritical.

The linear theory for a Witch profile gives a flux that is half that of a knife-edge in the deep ocean case, as shown in figure 16a (see St. Laurent *et al.* [18] and Llewellyn Smith and Young [19]). Various authors (see, for example, Balmforth *et al.* [20] and Pétrélis *et al.* [21]) have also considered the energy flux for other topographies and different values of the slope parameter ϵ. They find results that do not differ significantly when taking different



(a) (b)

Figure 16: The energy flux ratio for (a) the knife compared to the Witch and (b) the steep step compared to the inverse tan with $\epsilon = 1$ against δ. *From St. Laurent et al.* [18].

values of ϵ. The importance of this is that, in the deep ocean case, the slope (and hence the precise shape) of the topography is not the significant factor: the depth of the topography is the more important parameter.

It is also possible to use a similar technique to that given above to consider flow over a top hat ridge. Figure 17a shows the horizontal velocity profile for a steep step, similar to that for a top hat ridge, and figure 17b compares the energy flux from the top hat ridge to that from the knife edge.

5.2.2 A steep step

This is the extension of §§5.1.4 to an infinitely steep slope and is shown in figure 15b. We use the same procedure as in the previous subsection, however some modifications are required. The lack of symmetry across $x = 0$ implies that the wavenumbers to the left and right are not the same, and also $a_n \neq b_n$. Hence we set

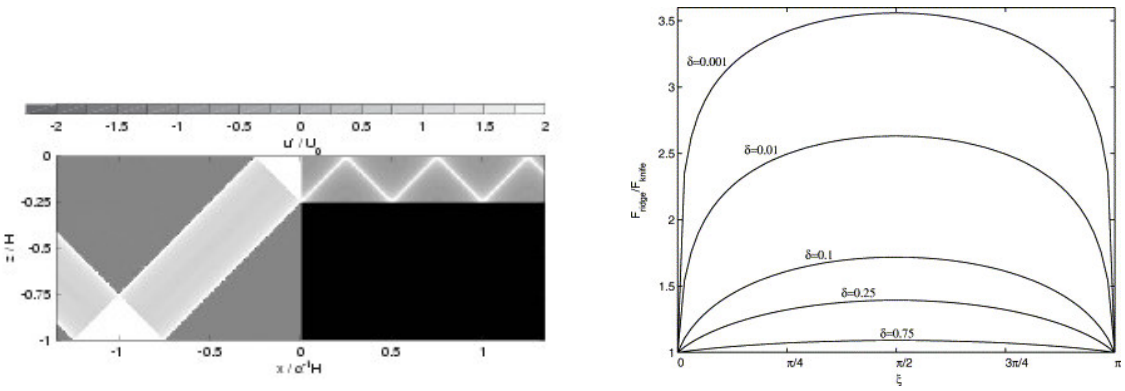
$$u_1 = U_0 \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi z}{H}\right) \cos(k_n x + \omega t), \quad w_1 = \alpha U_0 \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi z}{H}\right) \sin(k_n x + \omega t) \quad (27)$$

on the deep side and

$$\begin{aligned} u_2 &= U_0 \sum_{n=1}^{\infty} b_n \cos\left(\frac{n\pi z}{H - h_0}\right) \cos(-k_n' x + \omega t), \\ w_2 &= -\alpha U_0 \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi z}{H - h_0}\right) \sin(-k_n' x + \omega t) \end{aligned} \quad (28)$$

on the shallow side.

The boundary conditions are similar to (23)-(26). Noting that the barotropic tide, by continuity, must scale by $H/(H - h_0)$ to the right of the step, the boundary conditions



(a) (b)

Figure 17: (a) The baroclinic horizontal velocity profile resulting from a barotropic tide impinging on a steep step profile (the black rectangle). This profile is very similar to part of the profile seen for a top hat ridge. (b) Comparison between the energy flux of a top hat ridge with that of a knife edge. The fact that the curves become identical at $\alpha\pi L/H = \pi$, where L is the width of the ridge, is because the characteristic manages to fit into the ridge perfectly at this length. The worse agreement as δ become small is because the topography profile begins to look increasingly like two separate steps. *From St. Laurent et al.* [18].

become

$$\sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi z}{H}\right) = -\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi z}{H-h_0}\right), \qquad -H+h_0 \le z \le 0, \quad (29)$$

$$1+\sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi z}{H}\right) = \frac{H}{H-h_0} + \sum_{n=1}^{\infty} b_n \cos\left(\frac{n\pi z}{H-h_0}\right), \qquad -H+h_0 \le z \le 0, \quad (30)$$

$$1+\sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi z}{H}\right) = 0, \qquad -H \le z \le -H+h_0. \quad (31)$$

Multiplying (30) and (31) by $\cos(n\pi z/H)$ and (29) by $\sin[n\pi z/(H-h_0)]$ and integrating the equations vertically gives a matrix equation to solve for a_n and b_n of the form

$$a_m = \mathsf{A}_{mn} b_n + C_m, \qquad b_n = \mathsf{B}_{nl} a_l,$$

which may be solved numerically. Again, the flux may be found, and comparing it to the flux from the slope profile in §§5.1.4, we see, as for the ridge/knife edge comparison, that the difference is relatively small between supercritical and critical slope topography (see figure 16b).

Overall the conclusion is that, in the deep ocean, increasing the steepness of topographic features beyond the critical slope of the internal tide rays does not lead to a dramatic increase in energy flux into internal tides. However, increasing the height of the features, relative to the ocean depth, leads to a greater than quadratic increase in flux.

6 Open questions

There are many possible directions for research in internal tides:

1. Theoretical and numerical models for three-dimensional obstacle of finite slope and height.
2. The fate of low mode internal tides generated at bottom topography. Do they
   (a) Break down in the ocean interior?
   (b) Cascade to higher modes and turbulence on re-encounter with the sea floor?
   (c) Break as 'internal surf' on distant continental slopes?
3. Develop mixing parameterizations for use in ocean models.
4. Paleo-tides.

Notes by Josefina Arraut and Anja Slim.

References

[1] J. Pedlosky, *Waves in the Ocean and Atmosphere: Introduction to Wave Dynamics* (Springer, New York, 2003).

[2] D. E. Mowbray and B. S. H. Rarity, "A theoretical and experimental investigation of the phase configuration of internal waves of small amplitude in a density stratified fluid," J. Fluid Mech. **28**, 1 (1967).

[3] H. Sandstrom, Ph.D. thesis, University of California at San Diego, 1966.

[4] M. Rattray, J. G. Dworski, and P. E. Kovala, "Generation of long internal waves at the continental slope," Deep-Sea Res. **16 Suppl.**, 179 (1969).

[5] R. Regal and C. Wunsch, "M2 tidal currents in the western North Atlantic," Deep-Sea Res. **20**, 493 (1973).

[6] R. M. Hendry, "Observations of the semidiurnal internal tide in the western North Atlantic Ocean," Phil. Trans. R. Soc. **A286**, 1 (1977).

[7] W. H. Munk, F. E. Snodgrass, and M. Wimbush, "Tides offshore: transition from Califnornia coastal to deep sea waters," Geophys. Fluid Dyn. **1**, 161 (1970).

[8] R. D. Ray and G. T. Mitchum, "Surface manifestation of internal tides in the deep ocean: observations from altimetry and island gauges," Prog. Oceanogr. **40**, 135 (1997).

[9] J. M. Toole, J. R. Ledwell, K. L. Polzin, R. W. Schmitt, E. R. Montgomery, L. St. Laurent, and W. B. Owens, "The Brazil Basin Tracer Release Experiment," International WOCE Newsletter **28**, 25 (1997).

[10] R.-C. Lien and M. C. Gregg, "Observations of turbulence in a tidal beam and across a coastal ridge," J. Geophys. Res. **106**, 4575 (2001).

[11] C. S. Cox and H. Sandstrom, "Coupling of surface and internal waves in water of variable depth," Journal of the Oceanographic Society of Japan, 20th Anniversary Volume 499 (1962).

[12] T. H. Bell, "Topographically generated internal waves in the open ocean," J. Geophys. Res. **80**, 320 (1975).

[13] G. D. Egbert and R. D. Ray, "Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data," Nature **405**, 775 (2000).

[14] G. D. Egbert and R. D. Ray, "Estimate of M2 tidal energy dissipation from TOPEX/POSEIDON altimeter data," J. Geophys. Res. **106**, 22475 (2001).

[15] L. St. Laurent and C. Garrett, "The role of internal tides in mixing the deep ocean," J. Phys. Ocean. **32**, 2882 (2002).

[16] S. G. Llewellyn Smith and W. R. Young, "Conversion of the barotropic tide," J. Phys. Ocean. **32**, 1554 (2002).

[17] B. Sjöberg and A. Stigebrandt, "Computations of the geogrpahical distribution of the energy flux to mixing processes via internal tides and the associated vertical circulation in the ocean," Deep-Sea Res. **39**, 269 (1992).

[18] L. St. Laurent, S. Stringer, C. Garrett, and D. Perrault-Joncas, "The generation of internal tides at abrupt topography," Deep-Sea Res. **50**, 987 (2003).

[19] S. G. Llewellyn Smith and W. R. Young, "Tidal conversion at a very steep ridge," J. Fluid Mech **495**, 175 (2003).

[20] N. J. Balmforth, G. R. Ierley, and W. R. Young, "Tidal conversion by nearly critical topography.," J. Phys. Ocean. **32**, 2900 (2002).

[21] F. Pétrélis, S. G. Llewellyn Smith, and W. R. Young, "Tidal conversion at a submarine ridge," J. Phys. Ocean. (2004), *Submitted.*

Lecture 7: Tidal Bores

Chris Garrett

1 Introduction

In this lecture we discuss the formation of tidal bores in rivers. We discuss the effect of nonlinear terms in the shallow water equations. The convergence speed is obtained for tidal waves. This can be used to estimate the run length before a bore is formed. In §2 the nonlinear shallow water equations are discussed. We discuss a method of solution of the nonlinear shallow water equations in §3. A convergence speed is obtained in §4 and finally in §5 we discuss the Korteweg-de Vries Equation.

2 Nonlinear shallow water equations

The shallow water equations with nonlinear terms can be written as,

$$u_t + uu_x + g\zeta_x = 0, \tag{1}$$

$$\zeta_t + \frac{\partial[(h+\zeta)u]}{\partial x} = 0. \tag{2}$$

The effect of nonlinear term can be better understood using the following model equation,

$$u_t + (u+c)u_x = 0, \tag{3}$$

where c is constant. To obtain solution of (3) we can write, $u = \epsilon u_1 + \epsilon^2 u_2$. In the lowest order of ϵ we just obtain the simple wave equation,

$$u_{1t} + cu_{1x} = 0. \tag{4}$$

The solution of (4) is given as,

$$u_1 = F(x - ct). \tag{5}$$

If we consider equation at $O(\epsilon^2)$ then we get,

$$u_{2t} + cu_{2x} = -u_1 u_{1x}. \tag{6}$$

Consider an initial wave of sinusoidal form such that $u_1 = a\sin(k(x-ct))$. By substitution of u_1 in (6) we get,

$$u_{2t} + c\ u_{2x} = -\frac{1}{2}a^2 k \sin(2k(x-ct)), \tag{7}$$

with initial condition $u_2 = 0$ at $t = 0$. By solving (7) we get,

$$u_2 = -\frac{1}{2}ka^2 t \sin(2k(x - ct)). \tag{8}$$

When we are solving for u_2 in (7), this is a system of natural frequency c and is forced with a forcing frequency c. The value of u_2 is such that it increases linearly with time. This is counter intuitive because u_2 is $O(\epsilon^2)$ term and u_1 is $O(\epsilon)$ term, so u_1 should dominate over u_2, but this clearly cannot persist for long time.

An alternate interpretation of (3) can be in terms of its characteristics. The characteristic curves of (3) are given by the solutions of the ODEs,

$$\frac{dx}{dt} = u + c, \tag{9}$$

$$\frac{du}{dt} = 0. \tag{10}$$

That is, u is constant along each characteristic, which is the straight line,

$$x = x_0 + (u + c)t, \tag{11}$$

where x_0 is a constant labeling the various curves (by initial position). At $t = 0, u = a\sin(kx_0)$. Since this is a property of each characteristic, the solution is given implicitly by,

$$u = a\sin[k(x - ut - ct)]. \tag{12}$$

Solution (12) can be verified by substitution in (3). Evidently, the larger the amplitude, the faster the disturbance propagates with the result that the crests of the sinusoid travel faster than the troughs, which inexorably steepens the forward facing slopes. Eventually the crests catch the troughs and (12) predicts a multi-valued solution, which is unphysical. Instead, one must impose additional physics (such as dissipation) to regularize the problem, the outcome of which is typically the prediction of the formation of a shock.

3 Characteristics of the nonlinear shallow water equations

The nonlinear shallow water equations (1)-(2) can be rewritten taking $c^2 = g(h + \zeta)$ so that

$$u_t + uu_x + \frac{\partial c^2}{\partial x} = 0, \tag{13}$$

$$\frac{\partial c^2}{\partial t} + \frac{\partial (c^2 u)}{\partial x} = 0. \tag{14}$$

By adding and subtracting (13) and (14) we get,

$$\frac{\partial (u + 2c)}{\partial t} + (u + c)\frac{\partial (u + 2c)}{\partial x} = 0, \tag{15}$$

$$\frac{\partial (u - 2c)}{\partial t} + (u - c)\frac{\partial (u - 2c)}{\partial x} = 0. \tag{16}$$

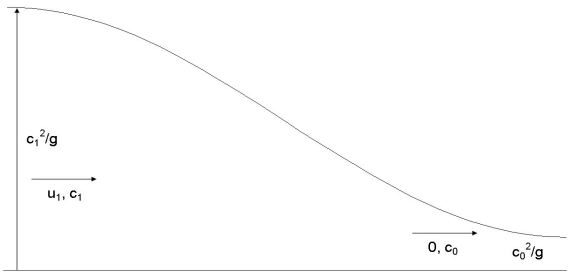


Figure 1: Steepening of tides causes formation of bores upstream in rivers. The distance upstream where bore forms can be calculated from the convergence speed.

The quantities $J_+ = u + 2c$ and $J_- = u - 2c$ are "Riemann invariants", each of which is constant along a particular characteristic curve on the (x,t) plane. J_+ is constant along the curves C_+ defined by,

$$\frac{dx}{dt} = u + c. \tag{17}$$

Similarly, J_- is constant along the curves C_- defined by,

$$\frac{dx}{dt} = u - c. \tag{18}$$

We can solve (13) and (14) by moving the Riemann invariants along the curves defined by C_+ and C_-. For most initial conditions, this procedure furnishes a simple numerical scheme. However, there are also special examples (such as the classic dam-breaking problem) that can be dealt with analytically.

4 Convergence Speed

We start by assuming that $J_- = B$ (constant everywhere) and J_+ is constant along the curves of C_+. For the rear and front of the wave we can write (see figure 1)

$$u_1 - 2c_1 = B, \tag{19}$$

$$-2c_0 = B. \tag{20}$$

From the equations above we can then write

$$u_1 = 2(c_1 - c_0). \tag{21}$$

The rear of the tidal wave travels at speed $u_1 + c_1$ and the front travels at speed c_0. So the convergence speed $u_{converg}$ is,

$$u_{converg} = 3(c_1 - c_0). \tag{22}$$

The steepening of the tidal wave occurs due to difference in velocity u and difference in heights c. Only one-third of the convergence speed is due to the difference in heights, the rest is due to the difference in velocity.

5 Korteweg-de Vries Equation

The Korteweg-de Vries equation can be found useful in a wide variety of wave problems in many diverse fields. In the ocean it can be found as a nonlinear representation of the shallow water equations. Thus, it can be used to model tidal waves. Its dimensional form is given as,

$$u_t + \left(\frac{3}{2}u + c_0\right)\frac{\partial u}{\partial x} + \frac{1}{6}c_0 h^2 \frac{\partial^3 u}{\partial x^3} = 0. \tag{23}$$

The 3/2 fraction in the non-linear term is to obtain a correct amount of steepening of the waves (based on the discussion in §4, one third of the steepening is due to convergence speed and the other two thirds is due to the difference in heights) and c_0 is the wave speed in shallow water. The last term in (23) is to make the equation consistent with the dispersion relation of the linear gravity waves given by,

$$\omega^2 = gk\tanh(kh) \approx c_0 k(1 - \frac{1}{6}k^2h^2). \tag{24}$$

The linearized shallow water approximation is useful only if the following two ratios are small:

$$\mu \equiv kh \ll 1, \qquad \epsilon \equiv \frac{A}{h} \ll 1. \tag{25}$$

Because of the severity of the the second restriction nonlinear theory of shallow water waves is needed. In order to derive this equation we first need to derive a low order Boussinesq-type equation. Since the velocity potential is analytic, we may expand it as a power series in the vertical coordinate. By substituting this series into the Laplace equation and by using the impermeable kinematic horizontal bottom boundary condition we can derive low order nonlinear shallow water equations [1].

$$\phi_{\mathrm{tt}} - \phi_{\mathrm{xx}} = \frac{\mu^2}{3}\phi_{\mathrm{xxtt}} - \epsilon\left(\phi_x^2 = +\frac{1}{2}\phi_t^2\right)_t. \tag{26}$$

Now let us assume a propagating wave in the positive x direction with a nondimensional wave speed of one and with a slow time variation: $\sigma = x - t, \tau = \epsilon t$. In terms of these variables, the derivatives become

$$\tfrac{\partial}{\partial x} \to \tfrac{\partial}{\partial \sigma}, \quad \tfrac{\partial}{\partial \sigma} \to -\tfrac{\partial}{\partial \sigma} + \epsilon\tfrac{\partial}{\partial \tau}. \tag{27}$$

By substituting these into the low order Boussinesq equation we get a nondimensional KDV equation in non-stationary coordinates

$$u_\tau + \frac{3}{2}uu_\sigma + \frac{\mu}{6\epsilon}u_{\sigma\sigma\sigma} = O(\epsilon, \mu^2). \tag{28}$$

Notes by Vineet K. Birman and Yaron Toledo.

References

[1] C. Mei, *The Applied Dynamics of the Ocean Surface Waves* (Teaneck, N.J. : World Scientific, Singapore, 1989).

Lecture 8: Tidal Rectification and Stokes Drift

Chris Garrett

1 Introduction

In this lecture, we examine how tidal flow which may be thought to be purely oscillatory can affect the mean flow. In §2 we examine flow separation and in §3 we look at the residual flow over a promontory. In §4 we describe Lagrangian vs. Eulerian motion. We see two arguments for how a clockwise mean flow around George's bank is generated by the tides.

2 Flow Separation

First consider the flow over a curved body (see Figure 1). Outside the boundary layer between points A and B, the body causes the streamlines to converge, resulting in acceleration of the outside flow, while the flow decelerates between B and C. Using Bernoulli's law, it can be shown that the accelerating region, the pressure decreases and in the decelerating region, the pressure increases. At the wall, the boundary layer equation is

$$\mu \frac{\partial^2 u}{\partial y^2} = \frac{\partial p}{\partial x}. \tag{1}$$

Thus where the pressure gradient is negative (a "favorable" gradient), $\partial^2 u/\partial y^2$ is also negative, leading to profile A. When the pressure gradient is positive (an "adverse" gradient), $\partial^2 u/\partial y^2$ is positive, leading to profile C. Using continuity, it is seen that the boundary layer grows in the the decelerating zone. The adverse gradient causes deceleration of the region immediately next to the body, as seen in the difference between profiles B and C. If the pressure gradient is sufficiently adverse, the flow can actually reverse at the body as seen in profile D. Where the reverse flow meets the forward flow, the flow will separate from the wall. Turbulent boundary layers are able to withstand the adverse pressure gradient better than a laminar boundary layer because the velocity profile above the body is more uniform and similar to the external flow than in the laminar case. Because of this, the separation point occurs farther along the body and consequently there is a thinner wake, which reduces the drag on the body. [1]

Flow separation can lead to tidal rectification. The next section discusses tidal flow around a promontory and the conditions for separation of tidal flow in the presence of friction.

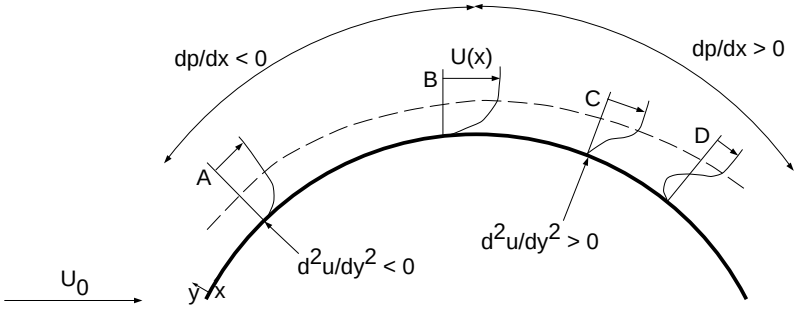


Figure 1: Flow past a curved body. The velocity profiles at various points along the body are shown along with the regions of adverse and favorable pressure gradients.

3 Residual circulation at a promontory

Strong tidal currents flow past the Gay Head promontory [2]. Measurements of the currents near the promontory show that there is a net flow away from the promontory. Figure 2 shows a schematic illustration of how this flow is created. When the tide flows past the promontory in one direction, the flow separates, creating an eddy. Then the tide reverses and does the same on the other side of the promontory. The combination of the induced currents creates a residual flow away from the promontory. Geyer and Signell (1990) measured this residual flow and found its average magnitude to be approximately 10% of the tidal flow magnitude. This phenomenon is discussed in the context of flow separation next.

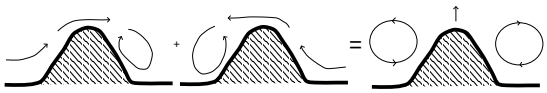


Figure 2: A residual current can be created by tidal flow around a promontory.

Outside the boundary layer next to the promontory, the following governing equation is valid:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = -g\frac{\partial \zeta}{\partial x} - \frac{C_d u\,|u|}{h}\ , \tag{2}$$

with x following the shape of the promontory. This flow is similar to that described in the previous section, although now friction is present. This has the effect of changing the direct relationship between the pressure gradient and acceleration or deceleration of the flow. Now, even for steady flow, the pressure gradient is balancing both the advective term and the friction term. While in the simple example the pressure gradient directly determined acceleration or deceleration, now the pressure gradient will continue to drive the flow in order to work against the friction. In this case, the flow will remain attached to the promontory and not separate if

$$\frac{C_D U^2}{H} > u\frac{\partial u}{\partial x},$$

where, U, H are the current speed and depth outside a narrow region, near the coast, over which the depth increases from 0 to H.

Consider the case of a small promontory (see Figure 3). Flow separation depends on if the pressure gradient is balancing the advective term or the friction term. If the advective term is larger than the friction term then flow separation will occur. The above condition for no separation can be scaled as

$$\frac{C_D U^2}{H} > \frac{U\Delta U}{a} \approx \frac{bU^2}{a^2}.$$

This condition can be restated as

$$\frac{a^2}{b} > \frac{H}{C_D}. \tag{3}$$

Geometry gives

$$a^2 + (R-b)^2 = R^2. \tag{4}$$

For the small promontory, $b \ll a$ so (4) can be rewritten as

$$\frac{a^2}{b} \approx 2R. \tag{5}$$

(3) now reduces to

$$R > \frac{H}{2C_D} \tag{6}$$

for the condition of no separation. For example, using typical values of $H = 20$ m and $C_D = 2.5 \times 10^{-3}$, R must be greater than 4 km to avoid flow separation. Separation is easier if $b \gg a$, or the promontory is much longer than it is wide.

For a given geometry there are controlling factors that dictate if the flow will separate:

$$\frac{b}{a}, \qquad Re_f = \frac{H}{C_D a}, \qquad K_c = \frac{U}{\omega a},$$

where Re_f is a friction Reynolds number, K_c is the Keulegan-Carpenter number, and ω is the tidal frequency. The ratio of the friction Reynolds number and the Keulegan-Carpenter number is

$$\frac{Re_f}{K_c} = \frac{H\omega}{C_D U} = \frac{\textit{spindown time}}{\textit{period}/2\pi}. \tag{7}$$

Tidal rectification occurs if the flow separates. The above analysis must be expanded for the more realistic situation of a varying depth H and 3D land forms to better evaluate where tidal rectification may occur.

4 Lagrangian Motion

There are two basic ways to describe a fluid flow. The Eulerian method, which is often a more natural description for observationalists, describes the flow of a fluid at certain fixed points. Velocity is measured as the flow past a particular point in space and time. The Lagrangian method refers to the flow following a fluid particle. Velocity is measured at

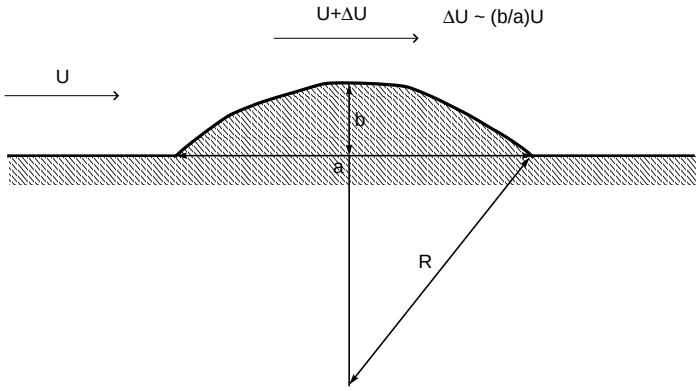


Figure 3: Flow past a small promontory. The promontory causes an increase in the velocity, ΔU as described in Section 2. For the analysis in the text to be valid, $b \ll a$.

the fluid particle, wherever it is in space and time. Using the subscripts E and L to mean Eulerian and Lagrangian, we have the following relationship between these two velocities for a particular particle at time t which started at position $\mathbf{x}_0$,

$$\mathbf{u}_L(\mathbf{x}_0, t) = \mathbf{u}_E\left(\mathbf{x}_0 + \int_0^t \mathbf{u}_L(\mathbf{x}_0, t')dt', t\right). \tag{8}$$

Note that the expression $\mathbf{x}_0 + \int_0^t \mathbf{u}_L(\mathbf{x}_0, t')dt'$ is simply the position $\mathbf{x}$ of the Lagrangian particle at time t.

If we take the average of these velocities in time and expand for small intervals in space, we arrive at the following relationship,

$$\langle \mathbf{u}_L(\mathbf{x}_0, t)\rangle = \langle \mathbf{u}_E(\mathbf{x}_0, t)\rangle + \left\langle \int_0^{t'} \mathbf{u}_L(\mathbf{x}_0, t')dt' \cdot \nabla \mathbf{u}_E(\mathbf{x}_0, t) \right\rangle + \dots \tag{9}$$

$$\approx \langle \mathbf{u}_E(\mathbf{x}_0, t)\rangle + \left\langle \int_0^{t'} \mathbf{u}_E(\mathbf{x}_0, t')dt' \cdot \nabla \mathbf{u}_E(\mathbf{x}_0, t) \right\rangle, \tag{10}$$

where $\langle\cdot\rangle$ is the mean. Assuming $\mathbf{u}_L \approx \mathbf{u}_E$, $\mathbf{u}_L$ in (9) can be replaced by equation $\mathbf{u}_E$ in (10). Dropping higher order terms, we have that the difference between the mean Lagrangian and Eulerian flow is the third term in (10), which is called Stokes drift.

The Stokes drift is like surfing. The more you stay with a wave, the more you drift forward; that is, you stay longer with the forward flow than if were standing still (Eulerian) in which case you would see the forward and backward flow for exactly the same amount of time. This forward drift is Stokes drift.

An interesting observation.

If we consider a plane wave, $\mathbf{u}_E = u_0 \cos(kx - \omega t)$ then the Stokes drift becomes

$$\left\langle \int_0^{t'} \mathbf{u}_E(\mathbf{x}_0, t')dt' \cdot \nabla \mathbf{u}_E \right\rangle = \frac{ku_0^2}{\omega} \left\langle \sin^2(kx - \omega t)\right\rangle = \frac{1}{2}\frac{u_0^2}{c},$$

where $c = \omega/k$ is the wave speed. In the non-rotating case, energy is equipartitioned between kinetic and potential energy. In this case,

$$\left\langle \int_0^{t'} \mathbf{u}_E(\mathbf{x}_0, t')dt' \cdot \nabla \mathbf{u}_E \right\rangle = \frac{1}{2}\frac{u_0^2}{c} = \frac{E}{c},$$

where E is energy. For tides, c can be quite large which means that the Stokes drift will be very small.

4.1 Example on a Slope: Georges Bank

If we consider a long wavelength tide approaching a bit of topography like Georges Bank in the Gulf of Maine, we can see the relationship between the Lagrangian and Eulerian flows, and the Stokes drift. In this argument, we consider the scale of the topography to be much smaller than the wavelength of the tide so that the wave is in phase all along the bank. Here we are considering an underwater sloping topography rather than a beach since at a beach the waves steepen and break, which are not the dynamics we're interested in. Lastly, the tidal flow makes ellipses due to the oscillation and the Coriolis force. As the wave encounters a sloping bottom, in order to conserve mass the tidal excursion in the horizontal must increase, thereby creating larger tidal ellipses in the shallow water on top of the bank.

4.1.1 Hand-Waving Argument

Kinematic argument. Using a kinematic argument, we can see that a particle which is advected towards the shallow water by the tide will see a larger tidal ellipse in shallow water, causing it to drift further to the bottom of the figure than it drifts back in the smaller tidal ellipse in deep water. See Figure 4. Thus, the Stokes drift is to the bottom of the figure (counter-clockwise around Georges Bank).

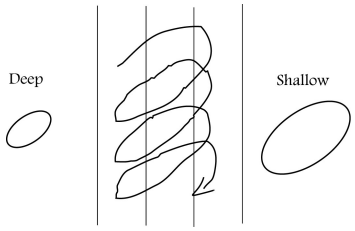


Figure 4: Kinematic argument: Stokes drift of a particle. This figure is a planar view of the underwater topography which is deep on the west and shallow in the east. The parallel vertical lines indicate bottom contours, and the closed ellipses are the tidal ellipses of the M_2 tide in deep and shallow water, with a gradient in size of the ellipse between. The curly line over the sloped topography is the path of a particle starting in the deep seas and experiencing the tidal ellipses as it moves to the east.

Dynamic consideration. From earlier considerations of drag, we know that it scales as $C_d u^2/h$, where C_d is the drag coefficient, u is velocity and h is depth. This means

that in shallow water, a particle would experience more drag per unit time than in deep water. If we assert that a particle must experience no net force and no net acceleration over a tidal cycle, then the particle must spend less time moving south in shallow water in shallow water than it does in deep water. See Figure 5. Thus Lagrangian flow must be to the north (clockwise around Georges Bank).

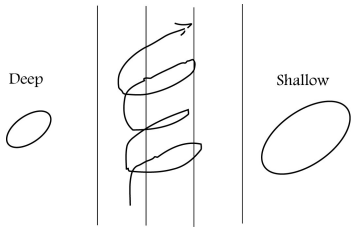


Figure 5: Dynamic consideration: Lagrangian flow of a particle in a tidal cycle. The tidal ellipses and topography are the same as in Figure 4, but the path of a particle using dynamic arguments is to the north rather than the south.

Eulerian flow. Since *Lagrangian* = *Eulerian* + *Stokes*, the Eulerian flow must be even stronger to the north (clockwise around Georges Bank) than the Lagrangian flow. Clockwise flow is indeed observed in the mean around Georges Bank.

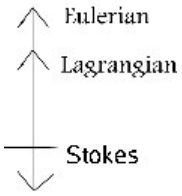


Figure 6: Eulerian flow around Georges Bank as the difference between Lagrangian flow and Stoke's drift.

4.1.2 Momentum Argument

This section will describe the tidal rectification around Georges Bank through continuity, Coriolis forces (changing the size of tidal ellipses) and bottom friction. If we consider the momentum equations of a tide encountering a bank averaged over a tidal period, we can solve for the resultant mean flow along the bank. Using 1 subscript to indicate tidal flow, the equations to solve are

$$\left\langle u_1 \frac{\partial u_1}{\partial x} \right\rangle - f \langle v \rangle + g \frac{\partial \langle \zeta \rangle}{\partial x} = 0, \tag{11}$$

$$\left\langle u_1 \frac{\partial v_1}{\partial x} \right\rangle = -\frac{\langle kv \rangle}{H}, \tag{12}$$

where $k = C_d(u^2 + v^2)^{1/2}$ is the frictional drag term at the bottom. We have assumed no mean across bank flow ($\langle u \rangle \approx 0$) and no pressure gradients along the bank ($\partial\zeta/\partial y = 0$).

Solving (12) gives the alongshore current for a given tide, then using v from the solution, (11) gives the across bank pressure gradient. Numerical solutions by Loder using deep and shallow water tidal ellipses and Georges Bank topography confirm the mean clockwise Eulerian flow [3]. He finds that for a tide of frequency ω with velocities u_1, v_1 the solutions have frequencies ω, 2ω and higher harmonics as well as the steady, rectified flow.

5 Conclusion

In this lecture we have seen how flow which is considered to be purely oscillatory from an Eulerian point of view can create a mean flow, either through interaction with topography or through Stokes drift. The argument for how Stokes drift results from the difference between Eulerian and Lagrangian velocity will be concluded in tomorrow's lecture with a vorticity description.

Notes by Danielle Wain and Eleanor Williams Frajka

References

[1] P. K. Kundu and I. M. Cohen, *Fluid Mechanics, 2ed* (Academic Press, New York, 2001).

[2] W. R. Geyer and R. Signell, "Measurements of tidal flow around a headland with a shipboard acoustic doppler current," J. Geophys. Res. **95(C3)**, 3189 (1990).

[3] J. W. Loder, "Topographic rectification of tidal currents on the sides of Georges Bank," J. Phys. Ocean. 1399 (1980).

Lecture 9: Tidal Rectification, Stratification and Mixing

Chris Garrett

1 Additional Notes on Tidal Rectification

This lecture continues the discussion of long-wavelength tidal flow over comparatively short topography such as George's Bank. The mean flow around the bank can also be understood via the following vorticity argument.

First, consider the vertical component of the fluid's relative vorticity,

$$\xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \tag{1}$$

ξ is defined as positive for counterclockwise flow, and negative for clockwise flow. Here, with water depth $H(x)$ and constant f, the evolution of the vorticity is given by

$$\frac{\partial \xi}{\partial t} + \mathbf{u} \cdot \nabla \xi - \frac{uf}{H^2}\frac{dH}{dx} = 0, \tag{2}$$

since the *potential vorticity* of the fluid is conserved following the flow:

$$\frac{D}{Dt}\left(\frac{\xi + f}{H + \zeta}\right) = 0. \tag{3}$$

We can see that changes in either the Coriolis parameter or the water depth will induce changes in the vorticity. If a fluid parcel is displaced north, its ambient vorticity increases ($df/dy > 0$), and it compensates by developing a negative relative vorticity, such that (3) is satisfied. Similarly, if a parcel moves into shallower water ($dH/dx < 0$) the relative vorticity has to decrease in order for potential vorticity to be conserved. Because of the analogy between planetary and topographical vorticity gradients, we can understand the mean flow due to a tidal current across the slope using the same argument that is made to explain the propagation of planetary Rossby waves.

For this analysis, we will neglect the nonlinear advective term in (2), and additionally assume that the rate of change of vorticity is a function of not just the Coriolis/depth term in (2), but that also of bottom drag. Defining λ as a friction parameter, we therefore model the rate of change of vorticity as

$$\frac{\partial \xi}{\partial t} = -\beta u - \lambda \xi, \tag{4}$$

where $\beta \equiv -\frac{f}{H^2}\frac{dH}{dx}$. If we plug in plane wave solutions

$$\xi = \xi_0 e^{-i\omega t}, \quad u = u_0 e^{-i\omega t},$$

then

$$\begin{aligned} -i\omega\xi &= -\beta u - \lambda\xi \\ \xi &= \frac{-\beta u}{\lambda - i\omega} \end{aligned} \tag{5}$$

Equation (5) helps us to interpret what happens to vorticity if there is both friction and a change of depth. In particular:

- If $\lambda = 0$, then the vorticity lags the current by a phase 90^o.
- If $\lambda \gg \omega$ then $\xi \propto -u$, so that vorticity and current are, at least in part, in phase.
- For any $\lambda \neq 0$, vorticity and current have opposite signs.

The importance of the drag term will be illustrated shortly.

Fluid that is displaced into shallow water takes on a negative relative vorticity, corresponding to clockwise rotation, which is then advected upslope. Fluid that is moving in the opposite direction takes on a positive relative vorticity, corresponding to counterclockwise rotation, which is advected downslope with that phase of the tide. Thus there is a downslope flux of positive vorticity. This flux is greatest over the steepest part of the slope, so its divergence leaves a residual vorticity, negative near the top of the slope and positive near the bottom. Since this generation process is independent of the position along the slope, the mean Eulerian current is parallel to the slope, and will flow with the shallow side to the right (when facing downstream). For a clockwise-rotating tidal current, the corresponding Stokes Drift will oppose it, which means that the net Lagrangian current is weaker than the Eulerian current.

Without friction, the current and vorticity are exactly out of phase, and the vorticity flux will therefore vanish when averaged over a tidal period, leaving zero residual current. Thus the presence of friction is necessary for a nonzero residual current.

2 Stratification and Mixing

Let us now examine stratification and mixing in a bay such as the Bay of Fundy [1]. Solar heating will, in principle, thermally stratify the water. This is what happens in a still lake, but in the oceans, where we have flow over bottom topography, the problem becomes more complex. We take three different approaches to the argument: first, since we're all geophysical fluid dynamicists here, we go through a simple dimensional argument and see what kind of solution is required. We then make a more mathematical argument by making conservation of energy considerations as in Simpson and Hunter (1974) [2]. Lastly, we use a highly simplified dynamical model to illustrate the role of mixing.

2.1 Dimensional Argument

First, we establish the dimensions of the variables of the problem, The tidal flow U has dimensions like LT^{-1}, where L is a length and T is a time. The depth of the fluid, H, has dimension L.

Since we are not so much interested in the heat flux, Q, as its effect on the water temperature, we instead consider the temperature change associated with the heat flux, or $Q(\rho c_p)^{-1}$. Actually, since temperature change also means a change in density, we will actually work with the quantity $\alpha Q(\rho c_p)^{-1}$, where $\alpha = \rho^{-1} d\rho/dT$. Then the buoyancy flux B is approximately given by

$$B = \frac{g\alpha Q}{\rho c_p}. \tag{6}$$

where B has dimensions L^2T^{-3}. The only dimensionless contribution of B, H, and U is BH/U^3, suggesting that a critical value of this separates *well-mixed* from *stratified* water.

2.2 Energy Argument

Suppose we model the warming of the bay by assuming the sun shines on the surface providing a heat flux of Q. Further suppose that only a small layer of height h is warmed by the heat giving the water in that layer a density of $\rho - \Delta\rho$. This warmer fluid lies above a larger layer of water with thickness H ($h \ll H$)and density ρ. See Figure 1 for a diagram of this model.

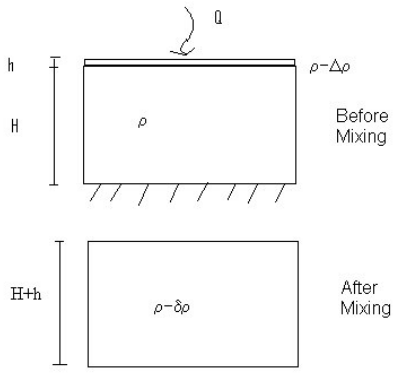


Figure 1: A Model of the Bay of Fundy

The change in temperature in the top layer can therefore be given by

$$h\frac{dT}{dt} = \frac{Q}{(\rho - \Delta\rho)\, c_p}, \tag{7}$$

where T is the temperature and c_p is the specific heat of water. Taking $\Delta\rho \ll \rho$, we can ignore the $\Delta\rho$ term in the denominator and express the rate of change of temperature in the top layer as

$$h\frac{dT}{dt} = \frac{Q}{\rho c_p}. \tag{8}$$

Using the chain rule, we can write

$$\frac{h}{\rho}\frac{d\rho}{dt} = \frac{h}{\rho}\frac{d\rho}{dT}\frac{dT}{dt} = -\frac{\alpha Q}{\rho c_p} = -\frac{B}{g}, \tag{9}$$

Now suppose that turbulent mixing homogenizes the fluid so that it all has density $\rho - \delta\rho$. Mass conservation requires

$$H\rho + h\left(\rho - \Delta\rho\right) = \left(H + h\right)\left(\rho - \delta\rho\right), \tag{10}$$

so that solving for $\rho - \delta\rho$ gives

$$\rho - \delta\rho = \frac{H\rho + h\left(\rho - \Delta\rho\right)}{H + h}. \tag{11}$$

The potential energy before mixing is given by

$$\frac{1}{2}gH^2\rho + ghH\left(\rho - \Delta\rho\right), \tag{12}$$

while after mixing the potential energy is

$$\frac{1}{2}g\left(H + h\right)^2\left(\rho - \delta\rho\right). \tag{13}$$

Therefore the change of potential energy is

$$\begin{aligned}
\Delta PE &= \frac{1}{2}g\left[\left(H + h\right)^2\left(\rho - \delta\rho\right) - 2hH\left(\rho - \Delta\rho\right) - H^2\rho\right] \\
&= \frac{1}{2}g\left[\left(H + h\right)\left(H\rho + h\left(\rho - \Delta\rho\right)\right) - 2hH\left(\rho - \Delta\rho\right) - H^2\rho\right] \\
&= \frac{1}{2}g\left[\left(H + h\right)^2\rho - \left(H + h\right)h\Delta\rho - \left(H + h\right)^2\rho + h^2\rho + 2hH\Delta\rho\right] \\
&= \frac{1}{2}g\left[hH\Delta\rho + h^2\left(\rho - \Delta\rho\right)\right].
\end{aligned} \tag{14}$$

Neglecting the quadratic terms in h gives us

$$\Delta PE = \frac{1}{2}gHh\Delta\rho. \tag{15}$$

Dividing (15) by Δt and taking the continuous limit we find that the rate of potential energy gain in the fluid is

$$\frac{dPE}{dt} = -\frac{1}{2}gHh\frac{d\rho}{dt} = \frac{1}{2}\rho BH, \tag{16}$$

where we have used (9) for the quantity $h\frac{d\rho}{dt}$.

If as in the dimensional analysis case, we assume that the tides are generating a flow in the horizontal direction of $u = U\cos\omega t$, the average rate of dissipation due to turbulence is

$$\begin{aligned}
c_d\rho < u^3 > &= c_d\rho\left(\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} U^3\cos^3\theta d\theta\right) \\
&= \frac{4}{3\pi}c_d\rho U^3.
\end{aligned} \tag{17}$$

Therefore, we can expect the lake to be well mixed if the rate of dissipation is sufficiently larger than the rate of potential energy increase needed to mix the water. Thus we expect that $B/(hU^3)$ must be less than a critical value, just as predicted by dimensional analysis.

2.3 A (Poor) Model

We again consider the bay to be a layer of water of height H being warmed by the sun which provides a buoyancy flux B at the surface. Let us define the buoyancy acceleration to be

$$b = -g\left(\frac{\rho - \rho_0}{\rho_0}\right). \tag{18}$$

The buoyancy frequency is defined by

$$N^2 = \frac{db}{dz}. \tag{19}$$

If we assume that the buoyancy change diffuses vertically through the bay, we can model the evolution of the buoyancy acceleration with the diffusion equation. That is

$$\frac{\partial b}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\partial b}{\partial z}\right), \tag{20}$$

where K is the diffusion coefficient. We must solve this equation with the boundary conditions that there is no buoyancy flux in at the bottom of the bay and that the flux at the top is B. Solving this problem gives

$$b = b_0 + \frac{Bt}{H} + \frac{Bz^2}{2HK}. \tag{21}$$

It remains to represent the diffusion coefficient, K, in terms of the parameters of the problem. Suppose that

$$K = CUHF\left(R_i\right) \tag{22}$$

where C is a constant and F is an unknown function of the dimensionless Richardson number R_i which represents how much stabilizing stratification there is in the system compared with destabilizing shear:

$$R_i = \frac{N^2}{\left(\partial U/\partial z\right)^2}. \tag{23}$$

If we take $N\left(H/2\right)$ to be the characteristic buoyancy frequency of the system, we have

$$R_i = \frac{B/2K}{\left(U/2H\right)^2}. \tag{24}$$

Solving for K and setting it equal to the definition of K in (22), we have

$$K = \frac{2BH^2}{R_i U^2} = CUHF\left(R_i\right), \tag{25}$$

so that

$$\frac{BH}{U^3} = \frac{1}{2}CR_iF\left(R_i\right). \tag{26}$$

Taking

$$K = 2\times 10^{-3}UH\left(1+R_i\right)^{-7/4}, \tag{27}$$

from Bowden and Hamilton (1975) [3], so that $F(x) = (1+x)^{-7/4}$. The maximum of $xF(x)$ occurs at $x = 4/3$. Plugging this value into (26) gives

$$\frac{H}{U^3} \approx 3 \times 10^4 B^{-1}. \tag{28}$$

For values of $BH/U^3 > 3 \times 10^4$, no steady solutions are possible so that the stratification will increase with time. This value may thus separate *well-mixed* from *stratified* condition. Using $B \approx 7 \times 10^{-8} m^2/s^3$ gives

$$\frac{H}{U^3} \approx 5000. \tag{29}$$

However, experiment shows that the critical value of H/U^3 for mixing is approximately $70 m^{-2} s^3$, which means that this model does capture the physics accurately. From observation, the efficiency of tidal mixing is 2.6×10^{-3},which implies that tidal currents lose most of their energy to turbulent dissipation other than to increasing buoyancy flux.

Notes by Lisa Neef and Dave Vener.

References

[1] J. K. C. Garrett and D. Greenberg, "Tidal mixing versus thermal stratification in the bay of fundy and gulf of maine," Atm. Ocean **16**, 403 (1978).

[2] J. Simpson and J. Hunter, "Fronts in irish sea," Nature **250(5465)**, 404 (1974).

[3] K. Bowden and P.Hamilton, "Some experiments with a numerical model of circulation and mixing in a tidal estuary," Estuarine and Coastal Marine Science **3(3)**, 281 (1975).

Lecture 10: Tidal Power

Chris Garrett

1 Introduction

The maintenance and extension of our current standard of living will require the utilization of new energy sources. The current demand for oil cannot be sustained forever, and as scientists we should always try to keep such needs in mind. Oceanographers may be able to help meet society's demand for natural resources in some way.

Some suggestions include the oceans in a supportive manner. It may be possible, for example, to use tidal currents to cool nuclear plants, and a detailed knowledge of deep ocean flow structure could allow for the safe dispersion of nuclear waste. But we could also look to the ocean as a renewable energy resource. A significant amount of oceanic energy is transported to the coasts by surface waves, but about 100 km of coastline would need to be developed to produce 1000 MW, the average output of a large coal-fired or nuclear power plant. Strong offshore winds could also be used, and wind turbines have had some limited success in this area.

Another option is to take advantage of the tides. Winds and solar radiation provide the dominant energy inputs to the ocean, but the tides also provide a moderately strong and coherent forcing that we may be able to effectively exploit in some way. In this section, we first consider some of the ways to extract potential energy from the tides, using barrages across estuaries or tidal locks in shoreline basins. We then provide a more detailed analysis of tidal fences, where turbines are placed in a channel with strong tidal currents, and we consider whether such a system could be a reasonable power source.

2 Barrages and Tidal Locks

The traditional approach to tidal power is with barrages, where a dam is set up to trap water at high tide and then release it through turbines at low tide, exploiting the potential energy of the trapped water. So a barrage plant will maximize its energy in a location with a large surface area and maximal tidal amplitude. The following table shows some of the most tidally active regions of the world:

	Mean Tidal Amplitude (m)	Basin Area (km^2)
La Rance, France	4	17
Bay of Fundy, Canada	5.5	240
Annapolis Royal, Nova Scotia	3.2	6
Severn Estuary, Great Britain	4	420
Garolim Bay, South Korea	2.5	85

The only major barrage plant is located at La Rance, Brittany, on the northern coast of France, which began operation in 1967. It has a capacity of 240 MW making it a significant energy source, given that typical coal and nuclear plants have an output of about 1000 MW. A disadvantage of this method is that the power source is clearly an intermittent one, which is why there is a compromise between releasing the energy at the lowest tide and at the period of greatest demand. La Rance was supposed to be the first of many tidal plants across France but their nuclear program was greatly expanded around this time, and so La Rance remains the only major tidal power plant in the world today. Although it has had an impact on the local ecology, it has been argued that the ecology was not particularly unique to begin with.

The only other commercially viable barrage plant is the Annapolis Royal station in Nova Scotia, which was activated in 1982 and has a capacity of about 20 MW. It was designed as an experimental station that would lead to an array of tidal power stations across the head of the Bay of Fundy, which is not only very large, but it also possesses the largest tides in the world. Such a network may be able produce an energy output as large as 5000 MW, but the environmental consequences could be severe; the more dire of predictions suggest that it could lead to a 5% drop in the Bay of Fundy tidal levels and a 10% rise in New England's levels. In any case, a better understanding of the effects of such a network would be needed before such an ambitious project can be attempted.

Other barrage plants exist, such as in Russia and China, but none have an output greater than about 500 kW. But the most dramatic of all proposed tidal plants is for a barrage across the Severn Estuary, between Wales and England. There is the possibility that over 8000 MW of power could be extracted, which could supply 12% of the current UK energy demand. Evaluations were conducted from 1974 to 1987, but the project was eventually shelved because of economic and environmental concerns. Other schemes, similar in size to the one of La Rance, may be implemented soon in South Korea.

3 Tidal Fences

Instead of directly using the variable potential energy of the oceans, it may be possible to utilize the strong tidal currents that are generated through narrow channels in certain parts of the world. We will attempt to quantify the actual amount of power that could be extracted from such a plant. This analysis has been done previously by Garrett and Cummins.[1]

We consider flow through a channel of variable cross-section (figure 1). The current speed $u(x,t)$ is assumed to be function of time t as well as x, but independent of the cross-channel position. The dynamical equation governing the flow is

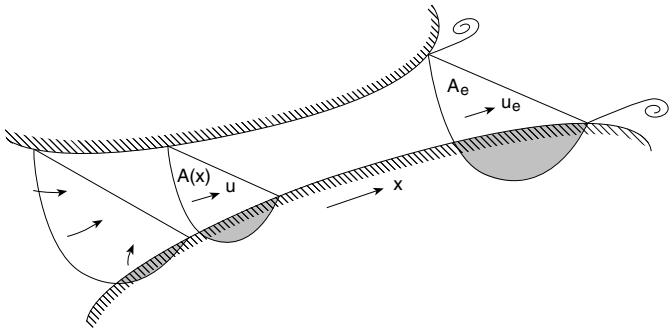


Figure 1: A channel connects two basins with different tidal elevations. At any instant, the current through the channel has speed u, a function of the local cross-sectional area $A(x)$, where x is the along-strait coordinate. At the channel entrance, the current enters from all directions, but at the exit it may separate as a jet with speed u_e, surrounded by comparatively stagnant water with elevation the same as that in the downstream basin.

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + g\frac{\partial \zeta}{\partial x} = -F \tag{1}$$

where the slope of the surface elevation ζ provides the pressure gradient to drive the flow and F represents an opposing force associated with natural friction and possibly the presence of turbines (in a uniform "fence" across the flow, a more efficient scheme than isolated turbines [2]).

If the channel is short compared with the wavelength of the tide (generally hundreds of kilometers even in shallow water), volume conservation implies that the flux along the channel $Au = Q(t)$, independent of x. (We neglect small changes in A associated with the rise and fall of the tide.) Using this in (1) and integrating along the channel implies

$$c\frac{dQ}{dt} - g\zeta_0 = -\int_0^L F\,dx - \frac{1}{2}u_e|u_e| \tag{2}$$

where $c = \int_0^L A^{-1}\,dx$ and $\zeta_0(t)$ is the sea level difference between the two basins, assuming that this difference is unaffected by the flow through the channel. We note that, if A increases rapidly at the ends of the channel, then the geometrical factor c is insensitive to their locations at $x = 0, L$. We allow for flow separation at the channel exit, with u_e denoting the current speed there.

We start by assuming that the natural frictional term and the head loss associated with separation at the exit are small, so that the natural regime has a balance between the sea level difference and acceleration. For $\zeta_0 = a\cos\omega t$, a sinusoidal tide with amplitude a and frequency ω, $Q = Q_0 \sin\omega t$ where $Q_0 = ga(\omega c)^{-1}$. The power generated at the turbines is the integral along the channel of the product of water density ρ, current u, cross-section A, and the local frictional force F representing the turbines. Thus the average power extracted from the flow by the turbines is

$$P = \overline{\int_0^L \rho F Q\,dx} = \rho\overline{Q\int_0^L F\,dx}, \tag{3}$$

with the over-bar indicating the average over a tidal cycle. This would need to be multiplied by a turbine efficiency factor to give the electrical power produced. We first

assume that the drag associated with the turbines is linearly proportional to the current. Then we may write $\int_0^L F\,dx = \lambda Q$ and $P = \lambda\rho\overline{Q^2}$, where λ is related to the number of turbines and their location along the channel. The governing equation

$$c\frac{dQ}{dt} - ga\cos\omega t = -\lambda Q \tag{4}$$

is easily solved to give

$$Q = \mathrm{Re}\left[\left(\frac{ga}{\lambda - ic\omega}\right)e^{-i\omega t}\right], \qquad P = \frac{\frac{1}{2}\rho\lambda g^2 a^2}{\lambda^2 + c^2\omega^2}. \tag{5}$$

As expected, P at first increases with λ (more power is generated as more turbines are added), but then decreases as too many turbines choke the flow. The maximum power, obtained when $\lambda = c\omega$, is

$$P_{\max} = \frac{1}{4}\rho g^2 a^2 (c\omega)^{-1} = \frac{1}{4}\rho c\omega Q_0^2. \tag{6}$$

The flow through the channel is then reduced to $2^{-1/2}$, or 71%, of its original value. We also note that this maximum power is independent of the location of the turbines along the channel (with the fewest being needed if they are deployed at the smallest cross-section). This result is independent of the representation of the turbine drag.

A more realistic representation of the turbine drag would be quadratic in the current speed. We have considered an arbitrary exponent, by non-dimensionalizing with $t = \omega^{-1}t'$, $Q = Q_0 Q'$, numerically solving

$$\frac{dQ'}{dt'} - \cos t' = -\lambda'|Q'|^{n-1}Q' \tag{7}$$

and evaluating $P/P_{\max} = 4\lambda'\overline{|Q'|^{n-1}Q'^2}$ as a function of λ' for different values of n (figure 2). The maximum power for $n = 2$ (quadratic drag) is very close to the value $P_{\max}$ derived for $n = 1$, and it does not seem that any other drag law could lead to more power.

We thus take $P_{\max}$ as a reasonable estimate of the maximum power available in the case of negligible background friction and exit separation effect. We may evaluate it in terms of conditions at the constriction with smallest cross-sectional area, $A_{\min}$, where the amplitude of the undisturbed tidal current is $u_{\max}$, if we write $c = \int_0^L A^{-1}\,dx = L_{\mathrm{eff}}/A_{\min}$, noting that this effective length L_{eff} is likely to be considerably less than the total channel length L. Then

$$P_{\max} = \frac{1}{4}\rho\omega L_{\mathrm{eff}} A_{\min} u_{\max}^2. \tag{8}$$

This may be compared with the reference value

$$P_0 = \frac{1}{2}\rho A_{\min}\overline{|u_{\max}\cos\omega t|^3} = \frac{2}{3\pi}\rho A_{\min} u_{\max}^3 \tag{9}$$

which gives the average undisturbed kinetic energy flux through the channel and is usually used as an estimate. The ratio is

$$\frac{P_{\max}}{P_0} = \frac{3\pi}{8}\frac{\omega L_{\mathrm{eff}}}{u_{\max}} \tag{10}$$

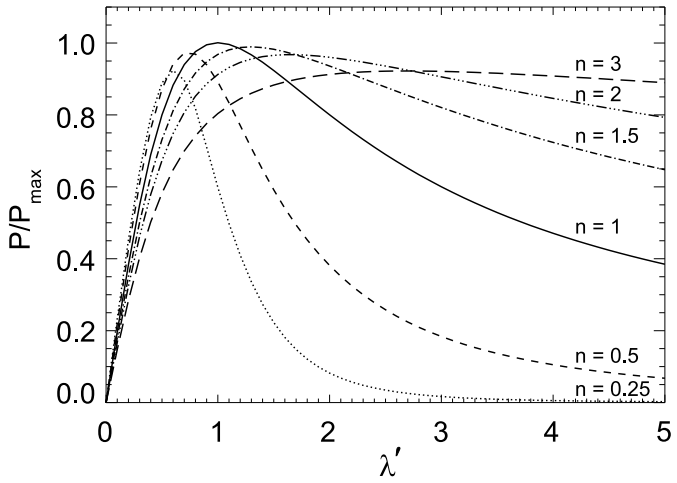


Figure 2: The scaled maximum power as a function of a parameter λ' representing the number of turbines in the channel, for various values of n, where the turbine drag is assumed proportional to the nth power of the current speed.

and can either be less than 1 (short channel, strong currents) or greater than 1 (vice versa). For example, with $\omega = 1.4 \times 10^{-4}\,\mathrm{s}^{-1}$, as for the semi-diurnal tide, $L_{\mathrm{eff}} = 5\,\mathrm{km}$, and $u_{\max} = 3\,\mathrm{m\,s}^{-1}$, then $P_{\max}/P_0 = 0.3$, so P_0 would overestimate the resource. With these values and $A_{\min} = 5 \times 10^5\,\mathrm{m}^2$ (as for a width of 5 km and a depth of 100 m), then $P_{\max} = 800\,\mathrm{MW}$.

The calculations so far have neglected background friction and separation effects. If these can be lumped together in a single term, as for quadratic friction in the channel, and represented by a value λ_0 in (7), we may still solve (7) but the power from the turbines is now $(\lambda' - \lambda_0)/\lambda'$ times the value calculated with $\lambda_0 = 0$. Here we examine the situations in which the acceleration is much less important than friction in the channel (as for a shallow channel) or the separation effect (as for a short channel). The momentum equation is then

$$-g\zeta_0 = -\int_0^L F_{\mathrm{turb}}\,dx - \alpha Q|Q|, \tag{11}$$

giving a balance at any instant between the surface slope, the drag associated $\int_0^L F_{\mathrm{turb}}\,dx$ with the turbines, and the internal friction and separation, where

$$\alpha = \int_0^L C_d (hA^2)^{-1}\,dx + \frac{1}{2}A_e^{-2}. \tag{12}$$

Here C_d is the drag coefficient, h the water depth, and A_e the cross-sectional area at the exit.

In the absence of turbines, the volume flux is $|Q_0| = (g|\zeta_0|/\alpha)^{1/2}$. With turbines, the instantaneous power is $\rho Q(g\zeta_0 - \alpha Q|Q|)$. This has a maximum of

$$P_{\max} = (2/3^{3/2})\rho\overline{Q_0 g\zeta_0} = 0.38 \times \rho\overline{Q_0 g\zeta_0} \tag{13}$$

where we have also averaged over a cycle. The flow at any instant is $3^{-1/2} = 0.58$ of what it was in the absence of the turbines, and 2/3 of the head loss along the channel is now associated with their operation.

If the friction within the channel is much less important than the separation effect, (13) may be written

$$P_{\max} = 0.38 \times \frac{1}{2}\rho A_e \overline{|u_{e0}|^3} \tag{14}$$

where $u_{e0} = (2g\zeta_0)^{1/2}$ is the exit speed at any instant in the natural regime. This formula is similar to (9), but we note the reduction factor of 0.38 and, particularly, the fact that it must be evaluated at the exit, likely giving a considerably lower value than would be obtained at the most constricted part of the channel. The turbines could still be deployed there, however.

In the preceding analysis we have assumed that the current is independent of the cross-sectional position. If there is shear vertically and across the channel, we may deal with the average speed in the acceleration term, but there will be minor changes in other parts of the theory. We have also ignored the back effect of changes in the channel flow on the forcing tides. This is likely to be small if the basins are large and deep, but there may be a positive feedback which will increase the head as turbines are introduced. This will slightly increase the available power.

Further consideration for particular sites will require more detailed analysis, but (8), (13) or (14), whichever is appropriate, will still give a good preliminary estimate of the power potential of a number of suggested sites around the world. We emphasize that the commonly used estimate, based on the energy flux in the natural state through the smallest cross-section, is incorrect and likely to be an overestimate. In general, significant power potential depends not only on the obvious factors of strong currents and large cross-section, but also on either the channel being long, or the currents at the exit separation, rather than just the constriction, being large.

Notes by Viswewaran Nageswaran and Marshall Ward.

References

[1] C. Garrett and P. Cummins, "The power potential of tidal currents in channels," Proc. Roy. Soc. A (2004), submitted.

[2] C. Garrett and P. Cummins, "Generating power from tidal currents," Journal of Waterway, Port, Coastal and Ocean Engineering **130**, 114 (2004).

Transmission of Rossby Wave Energy onto Gentle Slopes

Josefina Moraes Arraut

1 Introduction

We have investigated the transmission of energy of barotropic Rossby waves from a flat bottom region onto a gentle slope, on a β plane. We studied the simple case in which topography is a constant slope and both regions are large enough so that the plane wave approximation can be used. We have considered slopes gentle enough so that over one wavelength (L), the variation in depth ($\vec{\nabla H}\, L$) is small compared to the total depth H or, more precisely, $\frac{h}{H}$ is order Rossby number (ϵ). We also made use of the of the rigid lid approximation which renders the fluid horizontally non-divergent,

$$(Hu)_x + (Hv)_y = 0. \tag{1}$$

In our case, $H = H_0 - h$, where H_0 is the reference depth and h is the topographic elevation. A mass transport stream function can be defined:

$$Hu = -\psi_y, \qquad Hv = \psi_x, \tag{2}$$

which is used to write the linearized form of the potential vorticity equation:

$$\nabla_h \cdot \quad \frac{1}{H}\nabla_h \psi_t \quad + \quad \nabla_h \psi \times \nabla_h \quad \frac{f}{H} \quad \cdot \hat{z} = 0. \tag{3}$$

For our constant slopes this becomes:

$$\psi_{xxt} + \frac{\psi_{xt}}{H} h_x + \psi_{yyt} + \frac{\psi_{yt}}{H_0} h_y - \frac{f}{H} h_x \psi_y + \frac{f}{H} h_y \psi_x + \beta \psi_x = 0, \tag{4}$$

where h_x and h_y are constants.

In the above equation we wish to retain only terms with similar orders of magnitude. If we call X and Y the horizontal length scales of the wave in the x and y directions respectively, and T its time scale, we get for the equation above:

$$\frac{\psi}{X^2 T} + \frac{\psi}{X^2 T}\frac{X h_x}{H} + \frac{\psi}{Y^2 T} + \frac{\psi}{Y^2 T}\frac{Y h_y}{H} - \frac{f\psi}{YX}\frac{X h_x}{H} + \frac{f\psi}{XY}\frac{Y h_y}{H} \tag{5}$$

The second and fourth terms are small compared to the first and third ones and are disregarded. The last two terms are not disregarded because we are working in the low

frequency regime ($\omega << f$). We realized after the summer that this approximation is questionable and are still working on the consequences of removing it. We will however keep to it in this report for we still have a very limited understanding of the dynamics that would ensue otherwise.

With the approximation mentioned we get the equation:

$$\nabla_h^2 \psi_t + [\nabla\psi \times \nabla Q] \cdot \hat{z}, \tag{6}$$

where the gradient of ambient potential vorticity ($\nabla\vec{Q}$) is the constant vector: $\nabla\vec{Q} = \frac{f}{H_0} \quad h_x\,\hat{x} + \quad \frac{f}{H_0} \quad h_y + \beta \quad \hat{y}$.

If we choose the coordinate system so that the y-axis is aligned with $\nabla\vec{Q}$, (6) becomes

$$\nabla_h^2 \psi_t + |\nabla\vec{Q}|\psi_x = 0, , \tag{7}$$

which is the equation for planetary Rossby waves, with β substituted by $|\nabla\vec{Q}|$. Because h_x and h_y are constant, equations (6) and (7) have constant coefficients and admit plane wave solutions. Substituting

$$Ae^{i(kx+ly-\omega t)} \tag{8}$$

into (7), we get the dispersion relation:

$$\omega = \frac{-|\nabla\vec{Q}|k}{k^2 + l^2} \tag{9}$$

This expression shows that k and ω must have opposite signs. The k-component of the phase velocity, the quotient $\frac{\omega}{k}$, is always negative, so phase velocity is always contained in the half plane to the left of the $\nabla\vec{Q}$. When $|\nabla\vec{Q}|$ points to the North, as is the case for purely planetary Rossby waves, or for waves on a meridionally oriented slope growing towards North, phase propagation always has a westward component.

As noted by Longuet-Higgins, (9) can be written:

$$k - \quad \frac{-|\nabla\vec{Q}|}{2xs\omega}\Bigg)\Bigg)^2 + l^2 = \frac{|\nabla\vec{Q}|^2}{4\omega} \tag{10}$$

showing that the locus of points that satisfy this relation for fixed ω is a circle with centre $(\frac{-|\nabla\vec{Q}|}{2\omega}, 0)$, and radius $r = \frac{|\nabla Q|^2}{4\omega^2}$. If we choose the positive sign for ω we get the familiar illustration depicted in figure 1. We will make extensive use of this geometric interpretation, so we must look at it more closely.

First we notice that the vector defining the centre of the circle is $\frac{|\nabla\vec{Q}|}{2\omega}$, rotated by $\frac{\pi}{2}$ anti-clockwise. The centre vector and the radius having the same magnitude, the circle is tangent to the origin (this is a consequence of disregarding the displacement of the free surface, the geometrical effect of the free surface is to shorten the radius while leaving the centre vector unaltered). These circles are level curves of the function $\omega(k, l)$. The group velocity, which is the gradient in wave number space of this function:

$$\vec{c_g} = \quad \frac{\partial\omega}{\partial k}, \frac{\partial\omega}{\partial l} \quad = \quad \frac{|\vec{\nabla}Q|(k^2 - l^2)}{(k^2 + l^2)^2}, \frac{2|\nabla\vec{Q}|kl}{(k^2 + l^2)^2}\Bigg) \tag{11}$$

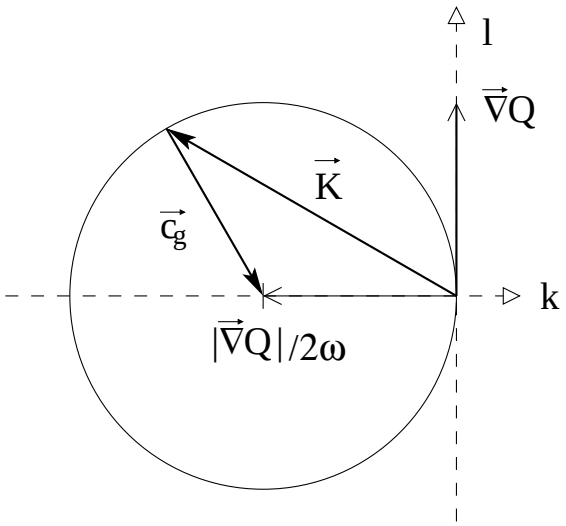


Figure 1: Longuet-Higgins' geometrical interpretation of the dispersion relation for Rossby waves on wave number space [1]. Points on the circumference are the values of k and l allowed by the dispersion relation. The vectors from the origin represent wave vectors and the ones from the origin to the centre give the direction of the group velocity.

must be radial. This becomes explicit if we rewrite (11) in polar coordinates:

$$(k, l) = K(\cos\alpha, \sin\alpha) \tag{12}$$

$$c_g = \frac{|\vec{\nabla}Q|}{K^2}(\cos 2\alpha, \sin 2\alpha) = \frac{|\vec{\nabla}Q|}{K^2}\vec{r}, \tag{13}$$

where α is the angle measured clockwise from the x-axis and $K^2 = k^2 + l^2$. We see that the group velocity grows with $|\vec{\nabla}Q|$ and decreases quadratically with the wave number. Figure (2) illustrates clearly why the group velocity is smaller for larger K^2.

We see from (13) that wave vectors lying on the right half of the circle have group velocities with a westward component, while those on the other half have group velocities with an eastward component and that the eastward moving waves are slower than the westward moving ones.

1.1 Energy flux

If we multiply equation (9) by ψ and rearrange terms, we get the equation of conservation of energy ([3])

$$\frac{\partial}{\partial t}\left[\frac{(\nabla\psi)^2}{2}\right] + \nabla\cdot\ \ -\psi\nabla\psi_t - \frac{1}{2}|\vec{\nabla Q}|\psi^2\hat{x}\ \ = 0. \tag{14}$$

The first term is the time variation of the kinetic energy, and the second is the divergence of a vector, called the energy flux vector ($\vec{S}$). By substituting the plane wave solution (8) into (14) and averaging over a wave period ($\langle\rangle$) we see that

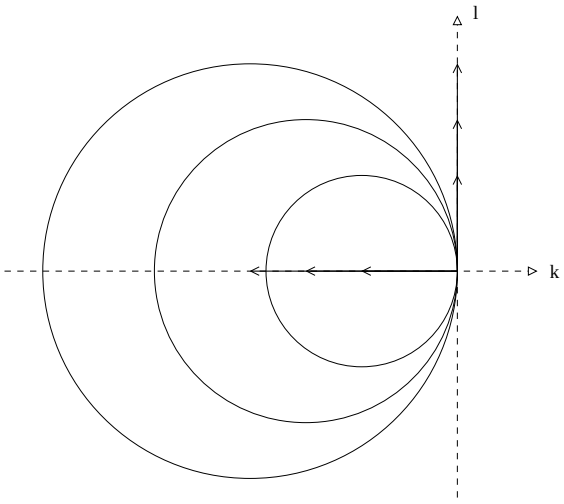


Figure 2: Level curves of $\omega(k, l)$.

$$\left\langle \vec{S} \right\rangle = \vec{c}_g \frac{A^2 K^2}{4} = \vec{c}_g \left\langle E \right\rangle, \tag{15}$$

the mean energy flux is the mean energy multiplied by the group velocity. The group velocity however is inversely proportional to the mean energy (11), so

$$\left\langle |\vec{S}| \right\rangle = A^2 |\vec{\nabla} Q|, \tag{16}$$

the magnitude of the energy flux for barotropic Rossby waves with a rigid lid is a function only of the gradient of ambient potential vorticity of the medium [1, 2].

We now proceed to formally solve the problem of energy transmission of Rossby waves from a flat bottom onto a straight slope. For clarity we will initially restrict ourselves to dealing with a slope in the zonal direction, an x-slope. In interpreting the formal solution to this and to the more general problem of a slope with arbitrary orientation, the geometrical interpretation of the dispersion relation reviewed above will become very useful.

2 Solution for the x-slope problem

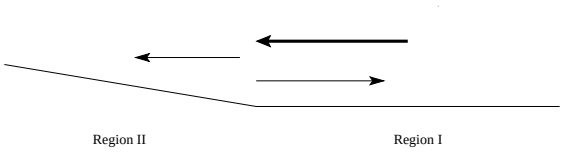


Figure 3: x-slope with impinging, reflected (region I) and transmitted (region II) waves

We want to find solutions to the equation:

$$\nabla^2\psi_t + \beta\psi_x = 0; \quad x > 0 \quad \text{region I} \tag{17}$$

$$\nabla^2\psi_t + h_x\psi_y + \beta\psi_x = 0; \quad x < 0 \quad \text{region II} \tag{18}$$

where $x = 0$ is the boundary between flat bottom and slope. We note that the boundary has been set perpendicular to $\vec{\nabla}h$. If the slope were to have a component along the boundary, there would be a step of variable height between the regions. We wish to restrict ourselves to the cases where $\vec{\nabla}h$ is discontinuous but not h, so our slopes will always be perpendicular to the boundary. The gradients of ambient potential vorticity for regions I and II are illustrated in figure 4. We remark again, for it will be relevant later, that they have the same along boundary component.

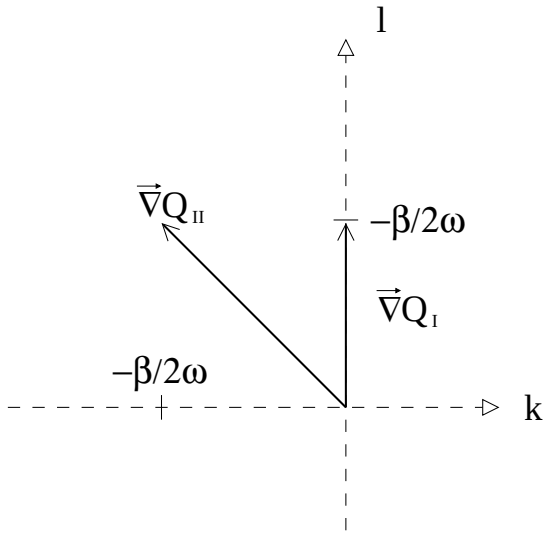


Figure 4: Ambient potential vorticity gradients for a flat bottom region and an x-slope with $f_0 h_x/H_0 = \beta$.

Equation (18) can be integrated trivially in the two open regions $x < 0$ and $x > 0$ because it admits plane wave solutions. The physical situation we want to represent is a plane wave impinging from region I onto region II (figure 3). In the stationary state we must have the motion in region I represented by the sum of two plane waves, one impinging and one reflected. The impinging wave must carry energy towards the boundary and the reflected wave away from it. In region II there should be only a transmitted wave. The kind of solution we seek is therefore:

$$\psi_i + \psi_r = Ae^{j(k_i x + l_i y - w_i t)} + Re^{j(k_r x + l_r y - w_r t)} \quad x > 0 \quad \text{region I} \tag{19}$$

$$\psi_t = Te^{j(k_t x + l_t y - w_t t)} \quad x < 0 \quad \text{region II} \tag{20}$$

where the subscripts i, r, t, stand for impinging, reflected and transmitted respectively and $j = \sqrt{-1}$. A, R, T are the constant amplitudes of the waves.

Only the parameters for the impinging wave are initially given. The others must be found by matching conditions at the boundary ($x = 0$). First of all, we require continuity of pressure and of mass transport at the boundary. If pressure were to be discontinuous, infinite accelerations would ensue. The second condition is conservation of mass since we cannot expect the boundary to be either a source or a sink of this quantity. These two conditions are degenerate in this case, because the flow is barotropic, and both lead to the expressions:

$$\psi_i + \psi_r = \psi_t\,, \quad x = 0 \tag{21}$$

$$A\,e^{j(l_i y - w_i t)} + R\,e^{j(l_r y - w_r t)} = T\,e^{j(l_t y - w_t t)} \tag{22}$$

For this to be true for all times and all $y's$, we must have:

$$l \equiv l_i = l_r = l_t \tag{23}$$

$$\omega \equiv \omega_i = \omega_r = \omega_t \tag{24}$$

$$A + R = T \tag{25}$$

We must also determine k_r and k_t and for this we have the two dispersion relations:

$$\omega = \frac{-\beta k_r}{k_r^2 + l^2}\,, \quad x > 0 \tag{26}$$

$$\omega = \frac{-\beta k_t + \frac{f}{H_0}\ h_x\, l}{k_t^2 + l^2}\,, \quad x < 0 \tag{27}$$

The relations give quadratic expressions for k, so radiation condition must be applied to decide upon the appropriate solution. That is, the value of k_r is chosen so that the reflected wave transmits energy eastward and the value of k_t is chosen so that the transmitted wave transports energy westward.

One more equation is needed to determine R and T. We obtain this by integrating (18) across $x = 0$ and taking the interval of integration to zero. In this way we solve the equation on the only point where it can't be trivially integrated. Our equation has a finite discontinuity in one of its, otherwise constant, parameters at $x = 0$. We want to know if this will impose a discontinuity in the across boundary derivative of the stream function, and if so, to quantify this jump.

$$\lim_{\epsilon \to 0} \quad \frac{\partial}{\partial t} \int_{\epsilon}^{-\epsilon} \quad \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \Bigg)\, dx - \int_{\epsilon}^{-\epsilon} Q_y \psi_x\, dx + \int_{\epsilon}^{-\epsilon} Q_x \psi_y\, dx \Bigg) = 0 \tag{28}$$

(22) and (23) show that all derivatives in the y-direction are continuous, so the second term in the first integral is zero and ψ_y can be removed from the third integral. Q_y is also continuous, so it can be removed from the second integral. This leaves us with the equation:

$$\lim_{\epsilon \to 0} \quad \frac{\partial}{\partial t} \int_{\epsilon}^{-\epsilon} \frac{\partial^2 \psi}{\partial x^2}\, dx - Q_y \int_{\epsilon}^{-\epsilon} \psi_x\, dx + \psi_y \int_{\epsilon}^{-\epsilon} Q_x\, dx \Bigg) = 0 \tag{29}$$

In our problem, Q_x has a finite discontinuity in $x = 0$ (going from β in region I to $\beta + \frac{f}{H_0} h_x$ in region II). Its integral is therefore continuous and the third term is zero. If we suppose a finite discontinuity in ψ_x, the second term is also zero and the first term quantifies this discontinuity. Since this is the only term left on the right, the discontinuity must be zero. ψ_x must therefore be continuous in our problem. From this condition we get the expression:

$$\lim_{x\to 0^+} (\psi_i + \psi_r) = \lim_{x\to 0^-} \psi_t \; k_i\psi_i + k_r\psi_r = k_t\psi_t|_{x=0}$$
$$k_i A + k_r R = k_t T \tag{30}$$

Substituting (25) into (30) we get the expression for the reflection and transmission coefficients in terms of the k wave numbers:

$$\frac{R}{A} = \frac{k_t - k_i}{k_r - k_t}, \qquad \frac{T}{A} = 1 + \frac{k_t - k_i}{k_r - k_t}. \tag{31}$$

The problem is now formally solved, for we have as many equations as unknowns. However, even for this simple case where we have fixed the orientation of the slope, the solutions depend on three parameters: ω, l, h_x, and it is hard to have a qualitative idea of their behavior.

We have, for instance, to worry about the existence of real solutions for k_t, given a value of l, in (26) and (27). A complex k_t would mean an evanescent wave in x, which would transport no energy. In other words, it would mean total reflection. Because the dispersion relations depend on both ω and h_x the condition for total reflection will depend on these two parameters. Furthermore, the reflection and transmission coefficients depend non-linearly on l, ω and h_x through k_r and k_t. Our main goal is to learn about the transmission of energy up the slope, it would be interesting to have some understanding of the role the different parameters play in this.

To gain a more intuitive grasp of the situation we will use the Longuet-Higgins circles. We have two regions with different $\vec{\nabla}Q$ so we will need to consider two circles at the same time. We will not be able to choose the coordinate system in which both dispersion relations have their simplest forms (9), so one of the circles will appear rotated from its familiar position shown in figure 1. Since the along boundary wave number is the same for the three waves, it is convenient to choose the direction of the boundary as the y-axis. In our case the direction of the boundary coincides with that of $\vec{\nabla}Q$ in region I. In figure 5 we see, for this choice of axis and $\frac{f}{H_0} h_x = \beta$, the Longuet-Higgins circles for regions I and II, on the left and right respectively.

As we noted earlier, the centre vectors of the circles are perpendicular to the direction of $\vec{\nabla}Q$ on the anti-clockwise sense, so all wave vectors (and phase velocities) lie to the left of $\vec{\nabla}Q$. The centres of both circles have the same k-coordinate, which is given by the along boundary component of $\vec{\nabla}Q$. The gradient is larger in region II, where there is the planetary β and the slope, so the circle is larger.

Figure 6 illustrates our problem. The value of l (determined by the choice of impinging wave) that all three waves must have in common is represented by the horizontal line segment. The wave vectors for the three waves must be the intersections of this segment with the circles. The direction of the group velocities for each wave are shown as the arrows

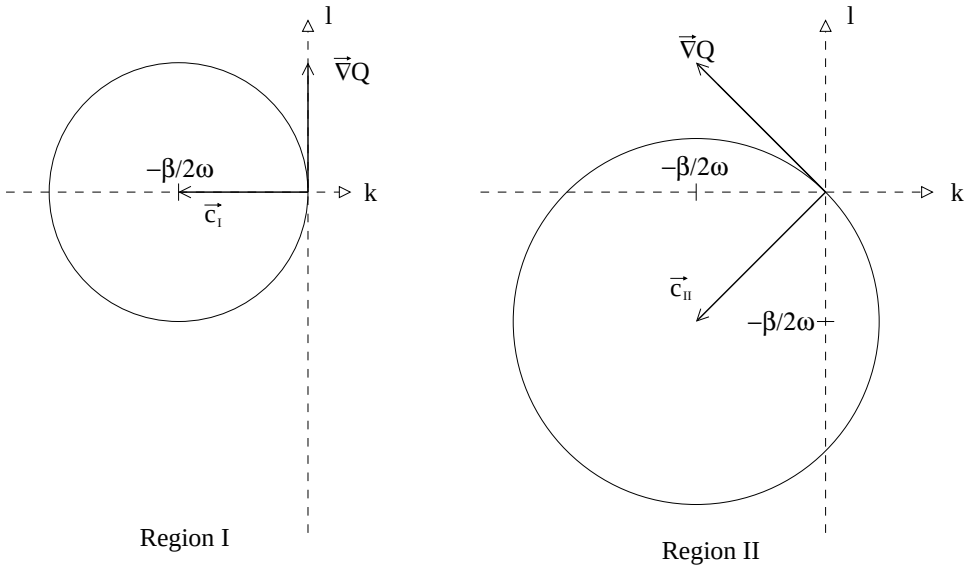


Figure 5: Longuet-Higgins circles for $\vec{\nabla Q}$ (region I) and $\vec{\nabla Q} = (\beta\,\hat{x},\ \frac{f}{H}\ h_x\,\hat{y})$ with $\frac{f}{H}\ h_x = \beta$ (region II).

from the circumference to the centre. Figure 5 shows how the reflected and transmitted wave vectors are chosen on the basis of the direction of the group velocity. We also notice that, upon crossing the boundary the group velocity veers to the south.

The issue of whether an impinging wave is able to propagate on to the slope is made clear in figure 5, where the two circles are drawn in the same set of axis. For the same frequency, the two media have different ranges of l which correspond to traveling waves. If an impinging wave has a value of l within the interval that is not allowed in region II, the wave will not be able to propagate on and there will be total reflection. All waves within the top shaded area cannot propagate up the slope. Likewise all the waves in the bottom shaded area would not be able to propagate from the slope to the flat bottom region.

The arrows in the figure show group velocities for the critical value of l, the value beyond which there is total reflection. We see that for this value the transmitted wave grazes the boundary. Beyond this value of l a transmitted wave would have to veer so much that it wouldn't even enter region II.

The fact that R and T depend on l through k and the importance of veering to this problem suggest that the angle of the group velocity is a more suitable variable than l.

A k wave number can be written in terms of this variable as

$$k = k_c - r\sin\gamma \tag{32}$$

where k_c, the k coordinate of the center, is Q_y, so

$$k = \frac{1}{2\omega}\ \ \vec{\nabla}Q\cdot\hat{y} - |\vec{\nabla}Q|\sin\gamma\ \ . \tag{33}$$

Q_y, the along boundary component of $\vec{\nabla Q}$, is the same for both media.

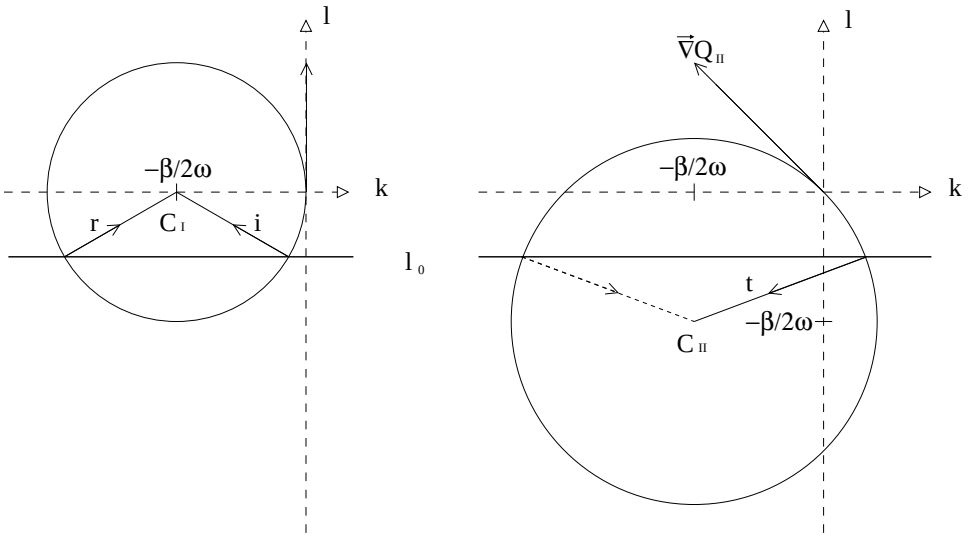


Figure 6: Longuet-Higgins circles for regions I (left) and II (right). The horizontal line segments represent the value of l which is common to all three waves. The intersections of this segment with the circles are the wave vectors that obey the dispersion relation for that value of l in each medium. The arrows from the circumference to the centre represent the group velocities for these vectors. For the impinging and transmitted waves the group velocities have a component to the west.

The coefficients R and T, having the form (wave number)/(wave number) are independent of ω. This is a consequence of the rigid lid approximation, which makes the centers of the circles and their radii tangent to the origin and the wave numbers proportional ω.

In these coordinates, R is written:

$$\frac{R}{A} = \frac{-|\vec{\nabla}Q_{II}|\sin\gamma_{II} + |\vec{\nabla}Q_I|\sin\gamma_I}{|\vec{\nabla}Q_I|\sin\gamma_I + |\vec{\nabla}Q_{II}|\sin\gamma_{II}} \tag{34}$$

Having said this and looking again at figure 7 we see that the critical value of l is actually better interpreted as a critical angle, or cutoff angle for transmission (γ_c). Remembering that the l coordinate of the centre of the circles is given by $\frac{Q_x}{2\omega}$ and the radius by $|\vec{\nabla}Q|$, it is easy to write (γ_c) in terms of the parameters of region II:

$$|\vec{r}| - \frac{Q_{IIx}}{2\omega} = |\vec{r}|\sin\gamma_c \tag{35}$$

$$\sin\gamma_c = 1 - \frac{Q_{IIx}}{|\vec{\nabla}Q_{II}|} \tag{36}$$

We now calculate $\frac{T(\gamma)}{A}$ and $\frac{R(\gamma)}{A}$ for $\frac{f}{H_0}\ \ h_x = \beta$.

The results shown in figure 8 are rather surprising. Not only does the transmission coefficient exceed 1, but it has a cusp at the cutoff angle, where we would expect energy transmission to be zero! All of this makes one worry about conservation of energy.

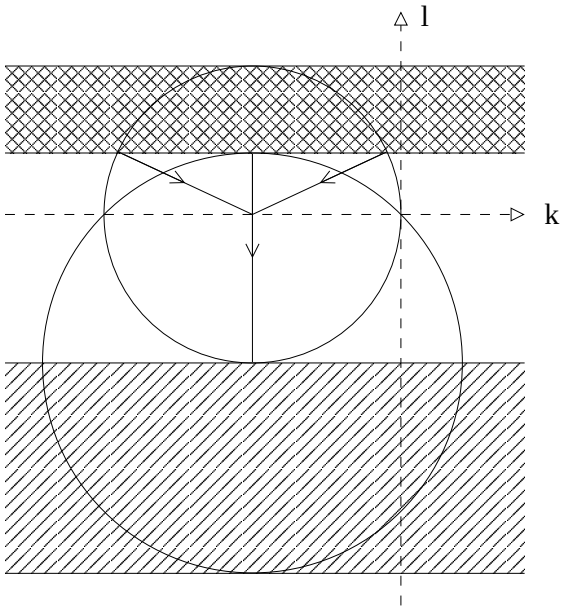


Figure 7: Total reflection: l wave numbers in the shaded areas are only allowed in one of the regions.

However, going back to (14), we see that it is the divergence of the energy flux that plays a part in energy conservation. We have no reason to expect accumulation of energy at the boundary, so we must have zero divergence of $\vec{S}$. As we see from (15) A, T and R give information about the magnitude of $\vec{S}$, but alone they don't contain enough information to account for its divergence.

For a plane wave, all fields are constant on a constant phase line, zero divergence at a point is the same as zero divergence on a finite interval on these lines. The condition for no accumulation of energy at the boundary is:

$$\int_{\Delta l} \vec{S}_I \cdot \hat{n}\, dl - \int_{\Delta l} \vec{S}_{II} \cdot \hat{n}\, dl = 0 \tag{37}$$

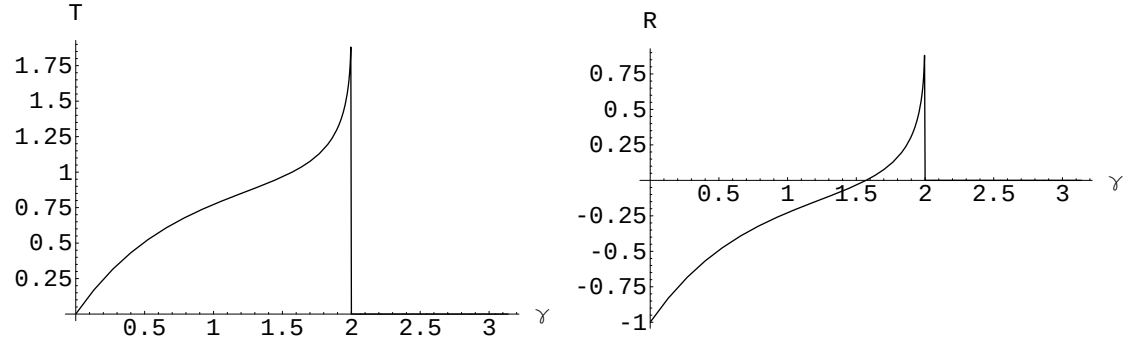


Figure 8: Transmission and reflection coefficients for a an x-slope with $Q_x = \frac{f}{H_0} h_x = \beta$.

We have seen that the group velocity veers when changing regions, so if we were to look at the energy flux vector across the boundary it would be something like figure 9 a).

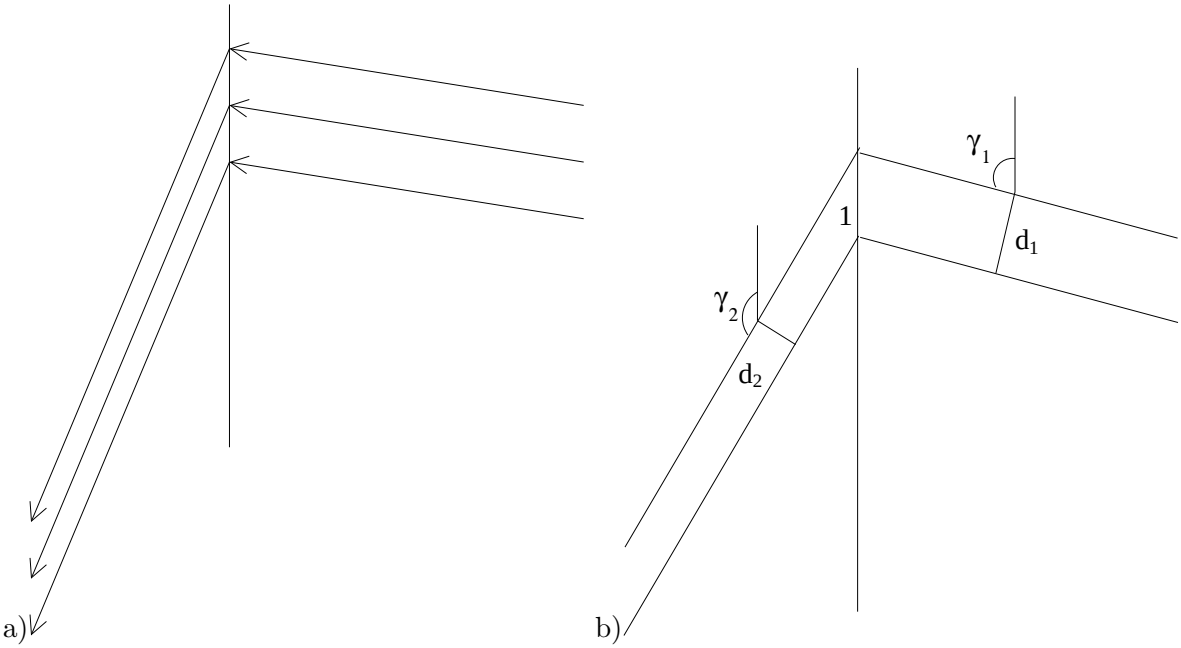


Figure 9: A wave beam, represented here by the energy flux vector, changes its direction and magnitude as it crosses the boundary in such away that its divergence is zero.

$\vec{S}$ (which is proportional to the group velocity) veers when crossing the boundary and the crossectional length of a beam is changed. For (37) to hold, the magnitude of $|\vec{S}|$ must change too. From (16) we see that $|\vec{S}|$ depends on $|\vec{\nabla}Q|$ and on the amplitude coefficient. As the veering depends on the angle of incidence and $|\vec{\nabla}Q|$ doesn't, all the angle dependence of $|\vec{S}|$ must be contained in the amplitude coefficient.

Figure 9 b) shows that the length factor for a beam which makes an angle γ with the vertical is $\sin\gamma$. With this last information we arrive at the condition for no accumulation of energy at the boundary:

$$A^2 - R^2 \quad |\vec{\nabla}Q_I| \sin\gamma_I = T^2 |\vec{\nabla}Q_{II}| \sin\gamma_{II} \tag{38}$$

or

$$\frac{T}{A}^{\ 2} \frac{\sin\gamma_I}{\sin\gamma_I} \frac{|\vec{\nabla}Q_I|}{|\vec{\nabla}Q_{II}|} + \frac{R}{A}^{\ 2} = 1 \tag{39}$$

We call the terms on the left the transmitted and reflected energy fractions. In figure 10 we show the result of evaluating the LHS of (39) for all angles of incidence. Energy is in fact conserved despite the awkward behavior of the amplitude coefficients seen in figure 8.

We are now ready to consider energy transmission onto the slope. Figure 11 shows the transmitted and reflected energy fractions as a function of the angle of incidence γ_I. We see that transmission does in fact go to zero at the critical angle (2 for this case, as

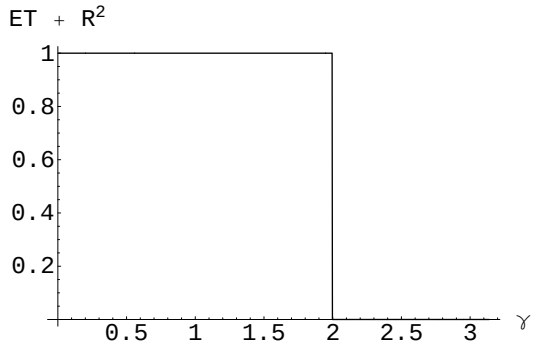


Figure 10: Sum of transmitted and reflected energy fractions calculated according to (39)

shown by the plots and as calculated from (36). We see also that energy transmission is maximum for $\gamma_I = \frac{\pi}{2}$. This corresponds to $\vec{K} = (0, 0)$, or a steady current. All energy from a steady current is transmitted onto the slope. The veering is such that the steady current remains perpendicular to $\vec{\nabla}Q$ in the new medium and the length factor for this veering exactly compensates for the change in $|\vec{\nabla}Q|$, in equation (39). Transmission falls off as $|\vec{K}|$ increases towards both sides on the circle.

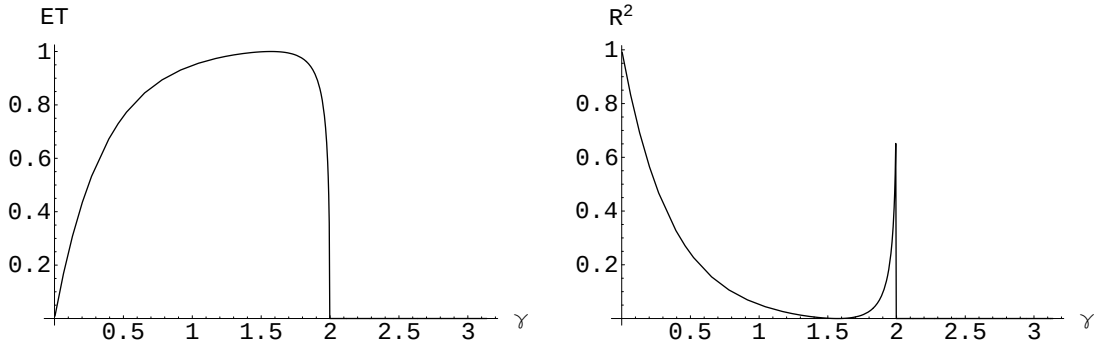


Figure 11: Transmitted and reflected energy fractions for the x-slope with $Q_x = \frac{f}{H_0} h_x = \beta$.

2.1 X-slopes of Different Steepness

Although we have up to now worked with a fixed value of h_x, changing it brings no surprises. As h_x grows, $\vec{\nabla}Q_{II}$ gets closer to the x-axis and the circle for region II closer to being tangent to the x-axis. This means that only impinging waves with a southerly group velocity component are allowed on the slope (cutoff angle $= \frac{\pi}{2}$). However, large slopes violate the assumption that $\frac{h_x L}{H}$ is small, made to obtain our wave equation and we shouldn't really worry about them. Figure 12 a) shows plots of the transmitted energy fraction for $\frac{f}{H_0}\ h_x = \beta, 2\beta, ...$ We see that the behavior in all these cases is qualitatively similar but with a cutoff angle approaching $\frac{\pi}{2}$ and the transmitted energy fraction falling off more rapidly with increasing $|\vec{K}|$ for the larger slopes. This is just as one would expect, for it should be harder for energy to proceed up steeper slopes.

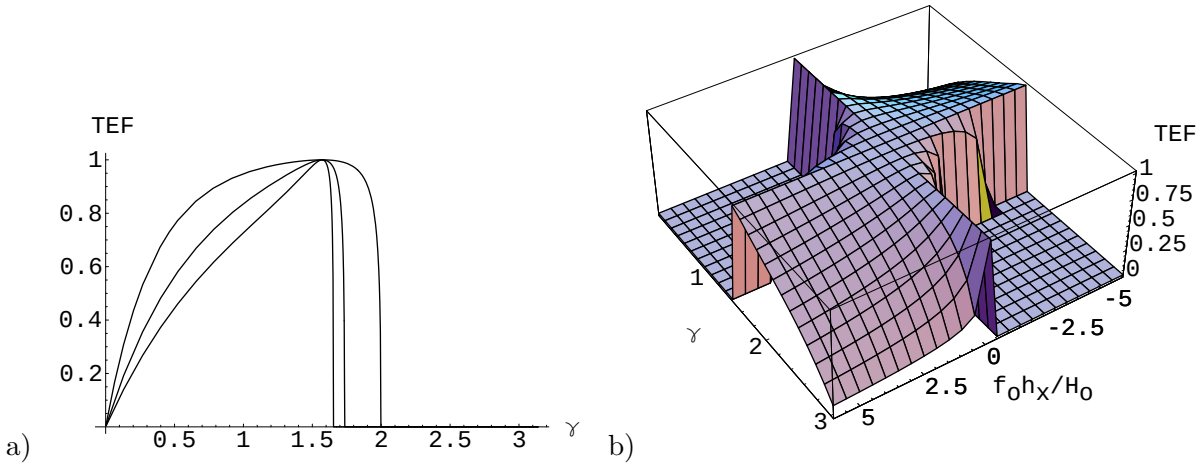


Figure 12: Transmitted energy fraction for a x-slope with $f_0 h_x/H_0$ varying from -6β to 6β and angle of incidence in relation to north (γ) varying from 0 to π.

3 Slopes with Different Orientations

Rotating the slope (and boundary) around brings more interesting consequences. By rotating the slope we can give the gradients of ambient potential vorticity in the different regions any relative orientation, including parallel and anti-parallel. In figure 13 a) we show the dispersion relations for a flat bottom region and a slope going up towards North. We see that all l wave numbers allowed in region I are also allowed in region II, so some energy can propagate up to the slope for all of wave numbers. As the relative angle between the gradients increases, their dispersion relations move apart in phase space and the two regions share less common l wave numbers. Transmitted energy fractions (TEF) for several of these cases are shown in figure 14 a) as a function of the angle of incidence measured from North ($\gamma + \theta$). We see that as the angle between the gradients increases, the band of waves allowed on the slope narrows, but the transmitted energy fraction is always 1 for $\gamma + \theta = \frac{\pi}{2}$ or incidence from the east, which corresponds to $(k = 0, l = 0)$. Figure 14 b) shows the transmitted energy fraction for all relative all orientations of the boundary and angles of incidence.

If a slope that grows towards the south is steep enough to overcompensate for the planetary β, the gradients of ambient potential vorticity for the two regions are anti-parallel. As can be seen in figure 13 a) , this means that they have no waves in common, except for $(k = 0, l = 0)$. The steady current would be the only possibility of Rossby wave energy exchange. However, this movement is East to West and therefore along the boundary. As far as Rossby waves go therefore, the two regions are isolated. Dispersion relations for this case are depicted in figure 13 b).

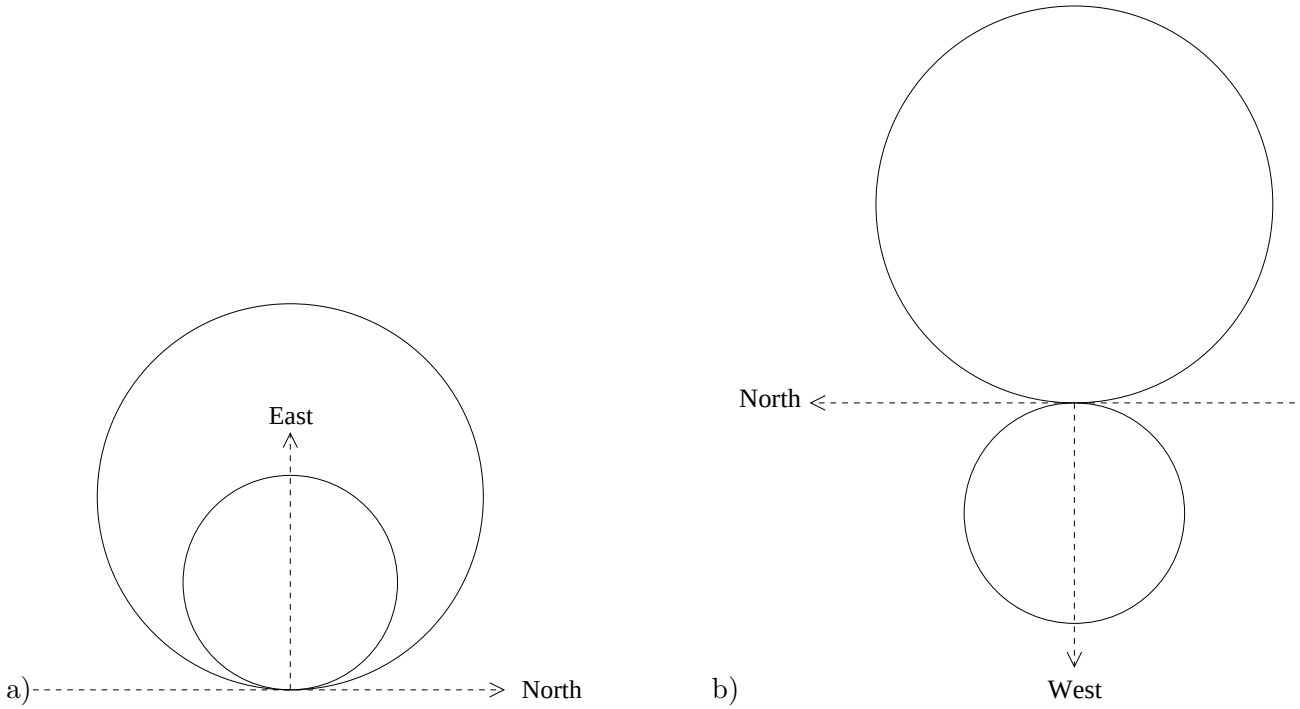


Figure 13: Dispersion relations for a flat bottom region and a) a y-slope growing towards north, b) a y-slope growing towards south.

4 Concluding Remarks

This work was intended as a simple first step in the investigation of energy transmission from mesoscale oceanic eddies onto the gentle continental slopes. In using the approximation of small depth perturbations we have confined ourselves to representing behavior not far from the foot of these slopes. In this context, the rigid lid approximation is not an additional restriction for it requires wavelengths to be small compared to the Rossby radius of deformation, which for the flat bottom open ocean is of order 2000Km.

We believe the main weakness of this work is the neglect terms 2 and 4 in equation 4. They are considered small when compared to the other slope terms (5 and 6) because, for the low frequency regime we are working on, $\frac{1}{T} << f_0$. However, terms with corresponding components of $\vec{\nabla} h$ have different spatial scales. For Rossby waves these spatial scales (X and Y) may be very different, invalidating the assumption on the relative sizes of the terms. We have realized this after the summer was over and are now working on understanding the effects of keeping the terms.

Acknowledgments

I would like to thank George Veronis for contributing his physical insight and experience to this project and for his kindness and patience in supervising it. Carl Wunsch for suggesting the problem and for the interesting conversations which allowed me a glimpse of his wide

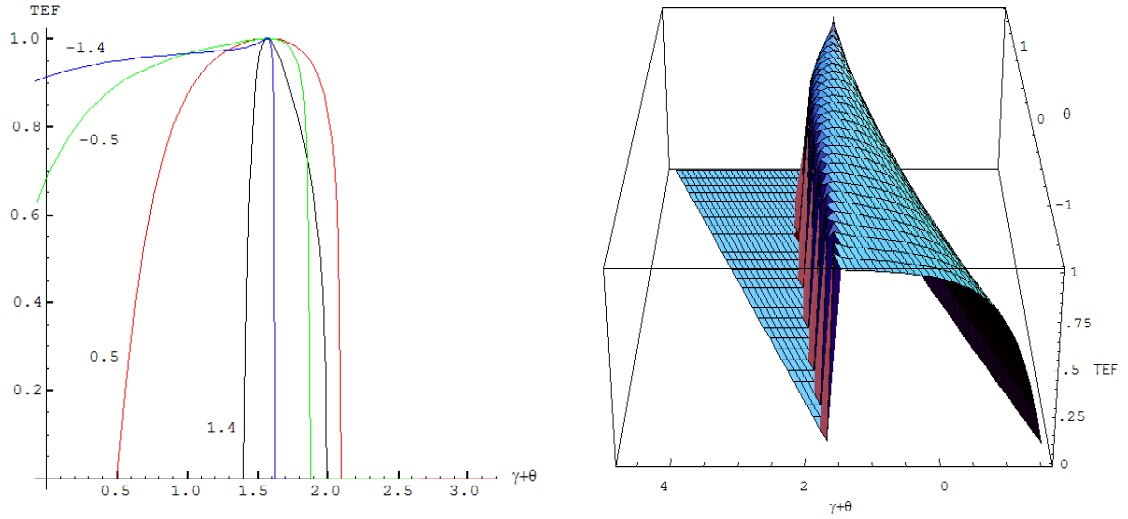


Figure 14: Transmitted energy fraction for $\vec{\nabla} h = -\beta$ for four different orientations of the boundary: 1.4, 0.5, -0.5, 1.4 radians with respect to North.

understanding of oceanography. Neil and Stefan for organizing the program for this year and, together with the other staff members providing such a rich working environment. I would like to thank also all the fellows and Shreyas and Shilpa for the friendship we cultivated during the summer.

References

[1] M.S. Longuet-Higgins, *Planetary waves on a rotating sphere*, Proc. Ray. Soc. A, **279** (1964) 446-473.

[2] M.S. Longuet-Higgins, *On group velocity and energy flux in planetary wave motions*, Deep-Sea Research, **11** (1964) 35

[3] J. Pedlosky, *Waves in the Ocean and Atmosphere*, Class Notes 12.802, MIT/WHOI Joint Program.

Forced Non-normal Convection

Vineet K. Birman

Abstract

We consider non-linear amplitude equations for double and triple diffusive convection close to marginal stability. We study the effect of non-normality on double and triple diffusive convection. We also consider the effect of small amplitude noise on the system. We observe that non-normality is not sufficient to take system away from basin of attraction of stable fixed points. Small amplitude noise along with non-normality can cause the system to go away from basin of attraction of stable fixed point. We also discuss the different regimes of behavior observed as the frequency of noise is changed.

1 Introduction

The importance of non-normality in transition to turbulence in shear flows has been investigated by many researchers. Non-normality was proposed as a possible explanation (c.f. Trefethen *et al.* (1993), Grossmann (2000)) for turbulence observed at Re numbers lower than those predicted by linear stability theory for shear flows. Non-normality is not important for Rayleigh convection as the linear operator is normal. We study the effect of non-normality in case of double diffusive and triple diffusive convection. Small noise can be amplified due to non-normality as observed by Gebhardt & Grossmann (1994). We also study the effect of small amplitude noise in case of double diffusive convection. In all cases we will study the amplitude equations for the system close to marginal stability.

Double diffusive convection, first suggested by Stommel *et al.* (1956) is a important geophysical phenomenon. Double diffusive convection is an important process in formation of stars. Double diffusive convection is also important in order to understand the ocean convection. It has been observed that close to marginal stability double diffusive convection can start even with small noise as shown in fig 1 from Shirtcliffe (1967).

Non-normality as a possible cause of instability was first considered by Trefethen *et al.* (1993) and Gebhardt & Grossmann (1994). Trefethen *et al.* (1993) suggested a nonlinear bootstrapping mechanism consisting of amplification of disturbances due to non-normality, resulting in nonlinear mixing such that output of nonlinear mixing is again amplified by non-normality. This mechanism is shown in fig 2. A more detailed description of works to study non-normality as possible pathway to turbulence is given in review paper by Grossmann (2000).

Most of the literature concerning study of non-normality is confined to their study for shear flows such as Baggett *et al.* (1993). The linear operator for shear flows are highly

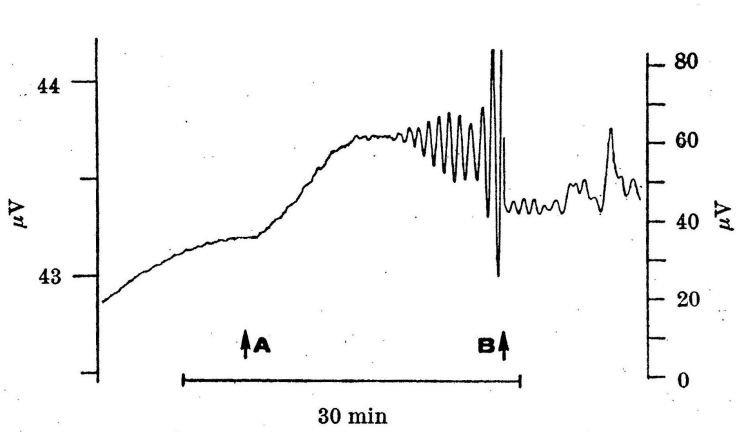


Figure 1: A section of temperature record near marginal stability. Differential output of two thermocouples is plotted against time. This transition is possible below marginal stability with small noise due to non-normality. Plot from Shirtcliffe (1967)

non-normal. The linear operator for the nonlinear amplitude equation obtained for double and triple diffusive convection obtained by Arneodo *et al.* (1985) and also Coullet & Spiegel (1983) is non-normal too. This suggests the possibility that non-normality can play important role even for convection flows.

Role of noise in non-linear dynamical systems has been studied for long time, Locher *et al.* (1998) is a review article listing important works done in the field. Even though we confine our study to single frequency small amplitude noise, our observations are in close correlation to observation for stochastic noise in dynamical systems. We find our results support the works done by Ganopolski & Rahmstorf (2002), where they suggest the possibility of climate change due to stochastic resonance.

In §2 we describe the system of equation considered. Linear analysis is described in §3. We obtain linear amplitude equations for our system in §4.We obtain nonlinear terms for the amplitude equations in §5. Finally the results are described in §6.

2 Equations

We consider two dimensional motion in a box of depth d (in z), width $\pi d/a$ (in x) and infinite length (in y), c.f. fig (3). We choose $\Delta\Theta$ and $\Delta\Sigma$, the magnitudes of the vertically impressed differences of temperature and salinity, as the units of temperature and salinity. The length and time units are d and d^2/κ_T, where κ_T is the thermal diffusivity. The equations are, c.f. Arneodo *et al.* (1985) & Chandrasekhar (1961),

$$\partial_t \Delta\Psi = \sigma\Delta^2\Psi - \sigma R\partial_x\Theta + \sigma\tau S\partial_x\Sigma + \sigma^2 T\partial_z\Upsilon + J(\Psi,\Delta\Psi), \quad (1)$$

$$\partial_t \Theta = \Delta\Theta - \partial_x\Psi + J(\Psi,\Theta), \quad (2)$$

$$\partial_t \Sigma = \tau\Delta\Sigma + \partial_x\Psi + J(\Psi,\Sigma), \quad (3)$$

$$\partial_t \Upsilon = \sigma\Delta\Upsilon - \partial_z\Psi + J(\Psi,\Upsilon). \quad (4)$$

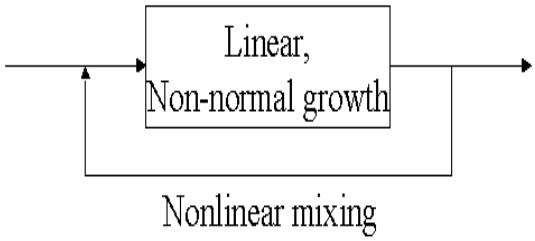


Figure 2: The suggested bootstrapping mechanism by Trefethen *et al.* (1993) leads to growth and instability.

where,
$J(f,g) = \partial_x f \partial_z g - \partial_z f \partial_x g, \quad \Delta = \partial_x^2 + \partial_z^2.$

The parameters are defined as,

$$\begin{aligned} R &= \frac{gd^3}{\kappa_T \nu} \left(\frac{\Delta\rho}{\rho}\right)_T, \\ S &= \frac{gd^3}{\kappa_S \nu} \left(\frac{\Delta\rho}{\rho}\right)_S, \\ T &= \left(\frac{2\Omega d^2}{\nu}\right)^2, \\ \sigma &= \frac{\nu}{\kappa_T}, \\ \tau &= \frac{\kappa_S}{\kappa_T}. \end{aligned} \tag{5}$$

where, R is thermal Rayleigh number, S is salinity Rayleigh number, T is Taylor number, σ is Prandtl number, τ is Lewis number, $(\Delta\rho/\rho)_T$ is the density difference solely due to imposed temperature difference and $(\Delta\rho/\rho)_S$ is density difference solely due to imposed salinity difference. Ω is the rotation rate about the z-axis.

The solution is independent of y. Ψ is the stream function for the x and z velocities, Θ and Σ are the deviations of temperature and salinity from their static values, and Υ is the y velocity divided by $\sigma T^{1/2}$.

We can express above equations concisely as

$$\partial_t \mathbf{L} U = \mathbf{M}_\lambda U + \mathbf{N}(U) \tag{6}$$

where

$$U(x,z,t) = \| \Psi, \theta, \Sigma, \Upsilon \|^T . \tag{7}$$

The parameter vector λ is

$$\lambda = (R, S, T, \sigma, \tau, a). \tag{8}$$

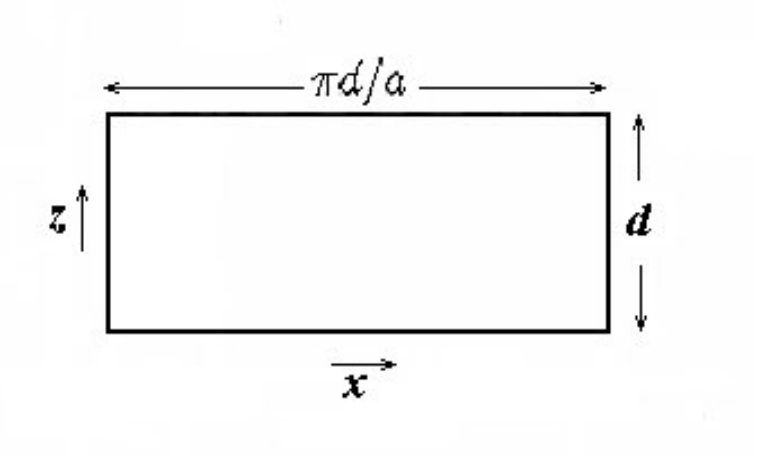


Figure 3: Two dimensional box of depth d and length $\pi d/a$.

The linear operators in eq. (6) are,

$$\mathbf{L} = \begin{vmatrix} \nabla^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{9}$$

and

$$\mathbf{M}_\lambda = \begin{vmatrix} \sigma\nabla^4 & -\sigma R\partial_x & \sigma\tau S\partial_x & \sigma^2 T\partial_z \\ -\partial_x & \nabla^2 & 0 & 0 \\ \partial_x & 0 & \tau\nabla^2 & 0 \\ -\partial_z & 0 & 0 & \sigma\nabla^2 \end{vmatrix} \tag{10}$$

The non-linear term is,

$$\mathbf{N}(\mathbf{U}) = (\mathbf{L}\partial_z U E')\partial_x U - (\mathbf{L}\partial_x U E')\partial_z U \tag{11}$$

where

$$E' = \begin{vmatrix} 1 & 0 & 0 & 0 \end{vmatrix}$$

is the transpose of E.

3 Linear Theory

The linearized form of equation (6) admits solution of form,

$$U = U_{mn} * \Xi_{mn} e^{st} \tag{12}$$

where, U_{mn} is a constant four component vector. Ξ_{mn} is defined as

$$\Xi_{mn} = \begin{vmatrix} \sin(max) & \sin(n\pi z) \\ \cos(max) & \sin(n\pi z) \\ \cos(max) & \sin(n\pi z) \\ \sin(max) & \cos(n\pi z) \end{vmatrix}$$

and the operator $*$ is defined such that

$$\begin{vmatrix} a \\ b \\ c \\ d \end{vmatrix} * \begin{vmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{vmatrix} = \begin{vmatrix} a\alpha \\ b\beta \\ c\gamma \\ d\delta \end{vmatrix}$$

s is a root of,

$$\mathbf{det} \parallel M_{mn} - L_{mn}s \parallel = 0 \tag{13}$$

where M_{mn} and L_{mn} are the matrix representations of the restrictions of $\mathbf{M}$ and $\mathbf{L}$ on Ξ_{mn}. They can be found by substitution of eq (12) into eq (6).

$$M_{mn}(\lambda) = \begin{vmatrix} \sigma q_{mn}^4 & ma\sigma R & -ma\sigma\tau S & -n\pi\sigma^2 T \\ -ma & -q_{mn}^2 & 0 & 0 \\ ma & 0 & -\tau q_{mn}^2 & 0 \\ -n\pi & 0 & 0 & -\sigma q_{mn}^2 \end{vmatrix} \tag{14}$$

and

$$L_{mn}(a) = \begin{vmatrix} -q_{mn}^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{15}$$

where

$$q_{mn}^2 = m^2a^2 + n^2\pi^2 \tag{16}$$

The fundamental in z is the most unstable of the vertical modes (as in Rayleigh- Bénard convection, Chandrasekhar (1961)), so we assume instability is encountered only for $n = 1$. When instability first occurs, it does so for a given value of a with $m = 1$ (Arneodo *et al.* (1985)). For $m = 1$ and $n = 1$, eq (13) is a quartic,

$$s^4 + P_3s^3 + P_2s^2 + P_1s + P_0 = 0 \tag{17}$$

where,

$$\begin{aligned} P_0 &= \pi^2\sigma^2\tau T q^2 - a^2\sigma^2\tau S q^2 - a^2\sigma^2\tau R q^2 + \sigma^2\tau q^8, \\ P_1 &= \pi^2\sigma^2(\tau+1)T - a^2\sigma\tau(\sigma+1)S - a^2\sigma(\sigma+\tau)R + \sigma(\sigma\tau+\sigma+2\tau)q^6, \\ P_2 &= \frac{\pi^2\sigma^2 T}{q^2} - \frac{a^2\sigma\tau S}{q^2} - \frac{a^2\sigma R}{q^2} + (\sigma^2+2\sigma\tau+2\sigma+\tau)q^4, \\ P_3 &= (1+2\sigma+\tau)q^2. \end{aligned} \tag{18}$$

where $q = q_{11}$.

We are looking for triple polycritiality point. This is a point in phase space of R, S and T such that $R = R_0$, $S = S_0$ and $T = T_0$ given by eq (23) where we have three marginally stable modes and all other modes are stable. Close to this point behavior of the system can be obtained by studying the behavior of these three stable modes. The condition for this is given by, c.f. Arneodo *et al.* (1985),

$$P_0 = P_1 = P_2 = 0 \tag{19}$$

and eq (17) becomes

$$s^3(s + P_3) = 0 \tag{20}$$

Dividing eq (17) by $(s + P_3)$ we get,

$$s^3 + P_2 s + (P_1 - P_2 P_3) + \frac{P_0 - P_3(P_1 - P_2 P_3)}{s + P_3} = 0 \tag{21}$$

In the neighborhood of the polycriticality condition (eq (19)), we have roots $|s| \ll P_3$ and we develop the remainder. We obtain an cubic critical polynomial for the marginally stable modes.

$$P_c(s) \equiv s^3 + \mu_2 s^2 + \mu_1 s + \mu_0 = 0 \tag{22}$$

where,

$$\mu_0 = \frac{P_0}{P_3}, \qquad \mu_1 = \frac{P_1}{P_3} - \frac{P_0}{P_3^2},$$
$$\mu_2 = \frac{P_2}{P_3} - \frac{P_1}{P_3^2} + \frac{P_0}{P_3^3}$$

The condition for polycriticality, eq (19), can be written as,

$$\begin{aligned} R_0 &= \frac{q^6(\tau + 2\sigma)}{a^2\sigma(1-\sigma)(1-\tau)}, \\ S_0 &= \frac{q^6\tau^2(1 + 2\sigma)}{a^2\sigma(\sigma-\tau)(1-\tau)}, \\ T_0 &= \frac{q^6\sigma(1 + \sigma + \tau)}{\pi^2(\sigma-\tau)(1-\sigma)}. \end{aligned} \tag{23}$$

Let λ_0 be a point on the polycritical surface defined by eq(23) or equivalently ($\mu \equiv (\mu_0, \mu_1, \mu_2) = 0$). Normal mode associated with $s = 0$ satisfies,

$$\mathbf{M}_{\lambda_0}\phi * \Xi_{11} = 0 \tag{24}$$

where ϕ is constant vector. Let $M_0 = M_{11}(\lambda_0)$ and $L_0 = L_{11}(a)$. For $s = 0$ in eq(13) is det $M_0 = 0$, and we solve for eigenvectors and generalized eigenvectors,

$$M_0\phi = 0, \quad M_0\psi = L_0\phi, \quad M_0\chi = L_0\psi. \tag{25}$$

Solving eq(25) we obtain,

$$\begin{aligned} \phi &= \| \, 1, \frac{-a}{q^2}, \frac{a}{\tau q^2}, \frac{-\pi}{\sigma q^2} \, \|^T, \\ \psi &= \| \, 0, \frac{a}{q^4}, \frac{-a}{\tau^2 q^4}, \frac{\pi}{\sigma^2 q^4} \, \|^T, \\ \chi &= \| \, 0, \frac{-a}{q^6}, \frac{a}{\tau^3 q^6}, \frac{-\pi}{\sigma^3 q^6} \, \|^T . \end{aligned} \tag{26}$$

4 Linear amplitude equation

The differential equations satisfied by coefficients (amplitudes) of the vectors are called amplitude equations. Considering the linearized form of eq(6),

$$\partial_t U = \mathbf{L^{-1}M}_\lambda U \tag{27}$$

Near polycriticality, only three characteristic solutions are nearly marginal, rest are damped normal modes. At polycriticality we define U as,

$$U = [A(t)\phi + B(t)\psi + C(t)\chi] * \Xi_{11}(x,z) \tag{28}$$

We want to derive an equation for amplitude vector $A = (A, B, C)$. For $\lambda = \lambda_0$ we have,

$$\dot{A} = JA \tag{29}$$

Extending it to parameter space where $\lambda \neq \lambda_0$. Suppose we select three vectors $\phi_\lambda * \Xi_{11}$, $\psi_\lambda * \Xi_{11}$ and $\chi_\lambda * \Xi_{11}$, that together form the stable modes. So, we have

$$U = [A(t)\phi_\lambda + B(t)\psi_\lambda + C(t)\chi_\lambda] * \Xi_{11}(x,z) \tag{30}$$

Thus the eq (29) is deformed into linear amplitude equation for $\lambda \neq \lambda_0$,

$$\dot{A} = K_\lambda A \tag{31}$$

where K_λ satisfies following conditions,

$$\begin{aligned} \mathbf{K}_{\lambda_0} &= \mathbf{J}, \\ \mathbf{det}(\mathbf{K}_\lambda - s\mathbf{I}) &= P_c(s). \end{aligned} \tag{32}$$

Since there are only three free parameters in $P_c(s)$, we want to express K_λ in only three parameters. The Jordan-Arnold canonical form has this property and we use it, c.f. Arneodo *et al.* (1985). This gives us a third order equation which can be written as,

$$\dddot{A} + \mu_2\ddot{A} + \mu_1\dot{A} + \mu_0 A = 0 \tag{33}$$

or writing it as,

$$\begin{aligned} \dot{A} &= B, \\ \dot{B} &= C, \\ \dot{C} &= -\mu_2 C - \mu_1 B - \mu_0 A. \end{aligned} \tag{34}$$

which is equivalently to

$$\dot{A} = J_\mu A \tag{35}$$

where,

$$J_\mu = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\mu_0 & -\mu_1 & -\mu_2 \end{vmatrix} \tag{36}$$

This is Jordan form of K_λ.

5 Nonlinear terms

We seek nonlinear terms in amplitude equation eq(35) of form,

$$\dot{A} = JA + g(A) \tag{37}$$

Employing the general method used in Coullet & Spiegel (1983), we express the time dependence of $U(x, z, t)$ in terms of $\mathbf{A}$ as,

$$U(x, z, t) = V\left[x, z, \mathbf{A}(t)\right] \tag{38}$$

Substituting eq(37) and eq(38) in eq(6) we obtain

$$\mathcal{L}\mathbf{V} = \mathbf{N}(\mathbf{V}) - (\mathbf{g} \cdot \partial_{\mathbf{A}})\,\mathbf{L}\mathbf{V} \tag{39}$$

where,

$$\begin{aligned} \mathcal{L} &= \mathbf{D}\mathbf{L} - \mathbf{M}, \\ \mathbf{D} = (\mathbf{J}\mathbf{A}) \cdot \partial_{\mathbf{A}} \quad &, \qquad \mathbf{M} = \mathbf{M}_{\boldsymbol{\lambda_0}}. \end{aligned}$$

We expand $\mathbf{V}$ and $\mathbf{g}$ in Taylor series in $\mathbf{A}$ and denote partial sum of all terms of degree κ as $\mathbf{V}_\kappa$ and $\mathbf{g}_\kappa$. Thus $\mathbf{V}_\kappa(x, z, c\mathbf{A}) = c^\kappa \mathbf{V}_\kappa(x, z, \mathbf{A})$. We now need to solve,

$$\mathcal{L}\mathbf{V}_\kappa = \mathcal{I}_\kappa - (\mathbf{g}_\kappa \cdot \partial_{\mathbf{A}})\,\mathbf{L}\mathbf{V}_\infty \tag{40}$$

where,

$$\begin{aligned} \mathcal{I}_\kappa &= \sum_{\alpha=1}^{\kappa-1} \mathcal{N}(\mathbf{V}_{\kappa-\alpha}, \mathbf{V}_\alpha) - \sum_{\alpha=1}^{\kappa-1} \mathbf{g}_{\kappa-\alpha+1} \cdot \partial_{\mathbf{A}} \mathbf{L}\mathbf{V}_\alpha, \\ \mathcal{N}(U, V) &= \mathbf{L}\partial_z V E' \;\partial_x U - \mathbf{L}\partial_x V E' \;\partial_z U. \end{aligned}$$

By successively solving for different orders we can obtain higher order terms. We will not do the entire solution here as it can be obtained from Arneodo *et al.* (1985). The final amplitude equation with non-linear terms can be written as,

$$\dddot{A} + (\mu_2 - k_3 A^2 - k_6 A\ddot{A})\ddot{A} + (\mu_1 - k_2 A^2 - k_4 A\dot{A} - k_5 \dot{A}^2)\dot{A} + (\mu_0 - k_1 A^2)A = 0 \tag{41}$$

where k_1,k_2,,k_6 are constant coefficients. Their expression can be obtained from Arneodo *et al.* (1985).
Since we want to consider the behavior of the system near the polycriticality point given by eq (23), we can consider asymptotic expansions of form,

$$\begin{aligned} R &= R_0 + \epsilon R_1 + \epsilon^2 R_2 + \ldots, \\ S &= S_0 + \epsilon S_1 + \epsilon^2 S_2 + \ldots, \\ T &= T_0 + \epsilon T_1 + \epsilon^2 T_2 + \ldots. \end{aligned}$$

and scaling given by,

$$\tau = \epsilon t, \qquad A = \epsilon^{3/2} x.$$

Expanding eq(41) using eq(42) and eq(42) to obtain amplitude equations of third order, which are

$$\dddot{x} + \hat{\mu}_2\ddot{x} + \hat{\mu}_1\dot{x} + \hat{\mu}_0 x = k_1 x^3 \tag{42}$$

We have x^5 term in the next order nonlinear terms. We include x^5 in some cases where we need additional fixed points for the system to prevent the solution from blowing up. So the amplitude equation if we include the x^5 term from the higher order nonlinear terms is given as,

$$\dddot{x} + \hat{\mu}_2\ddot{x} + \hat{\mu}_1\dot{x} + \hat{\mu}_0 x = k_1 x^3 - l x^5 \tag{43}$$

This amplitude equation can be rederived for double diffusive convection case where $T = 0$. The equation we will obtain (Coullet & Spiegel (1983)) will be (x^5 nonlinearity included),

$$\ddot{x} + \hat{\mu}_1\dot{x} + \hat{\mu}_0 x = k_1 x^3 - l x^5 \tag{44}$$

We intend to study the effect of the non-normal operator on this system looking for possibility of going out of basin of attraction of a stable fixed point of the system with small initial perturbations. We will also study the effect of noise on this system and possibility of chaos due to presence of small amplitude noise.

6 Results

The second order amplitude equation with A^5 nonlinearity (eq (44)) has five fixed points, given by

$$A = 0, \quad A = \pm\left[\frac{k_1 + \sqrt{k_1^2 - 4l\mu_0}}{2l}\right]^{1/2}, \quad A = \pm\left[\frac{k_1 - \sqrt{k_1^2 - 4l\mu_0}}{2l}\right]^{1/2} \tag{45}$$

The bifurcation diagram for these set of equations is shown in fig 4 for $\mu_1 = 0.71$, $k_1 = 1$ and $l = 0.5$. The solid lines indication the stable solution while the dashed lines indicate the unstable solution.

We have used 4th order Runge-Kutta Scheme to integrate the ODE's. We have used constant time step. The time step was chosen to be $\Delta t = 0.001$ for all simulations discussed here. This time step is small and integration converges for all cases discussed.

6.1 Basin of attraction of origin

Basin of attraction is defined as the region of states, in a dynamical system, around a particular stable steady state, that lead to trajectories going to the stable steady state. We are studying the basin of attraction of the origin.

Fig 5 shows the basin of attraction for eq (44) for the stable fixed point at origin. The basin of attraction is shaped like a slit. It does not close as shown in the figure but extends till infinity. As the figure was calculated numerically it gets truncated.

We have observed that due to non-normal nature of eq (44) even if we give a small initial perturbation inside the basin of attraction the amplitude grows and goes close to the boundary of basin of attraction before decaying back to origin. This suggested that in presence of small noise the system could go away from origin even with very small initial disturbances.

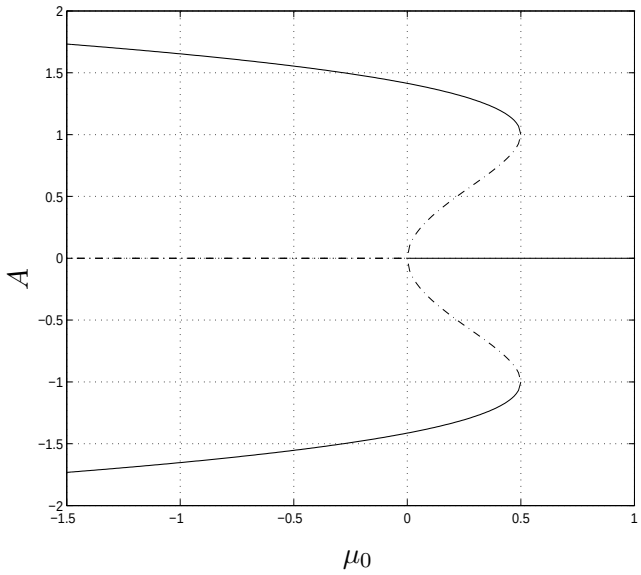


Figure 4: Bifurcation diagram for the amplitude equation. Solid line represents stable solution and dashed line is unstable solution.

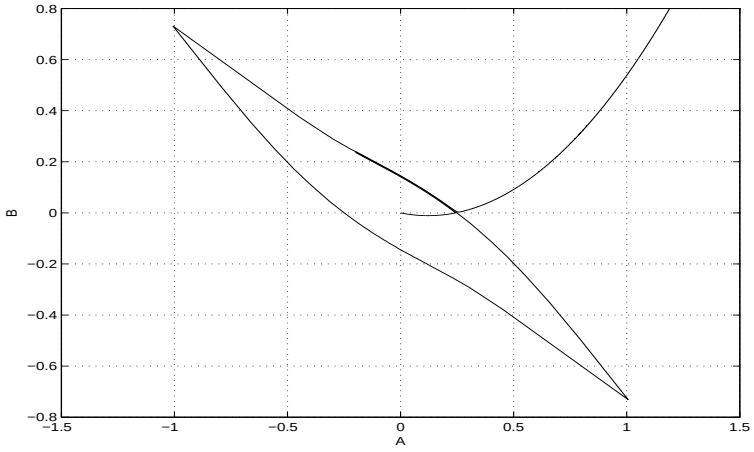


Figure 5: The basin of convergence of Origin for second order amplitude equation with $\mu_1 = 0.5$, $\mu_0 = 0.006$, $k_1 = 1$ and $l = 0$. Two trajectories are shown. One starts just inside the basin and ends up at the origin, while the second trajectory starts at a point just outside the basin and goes to infinity.

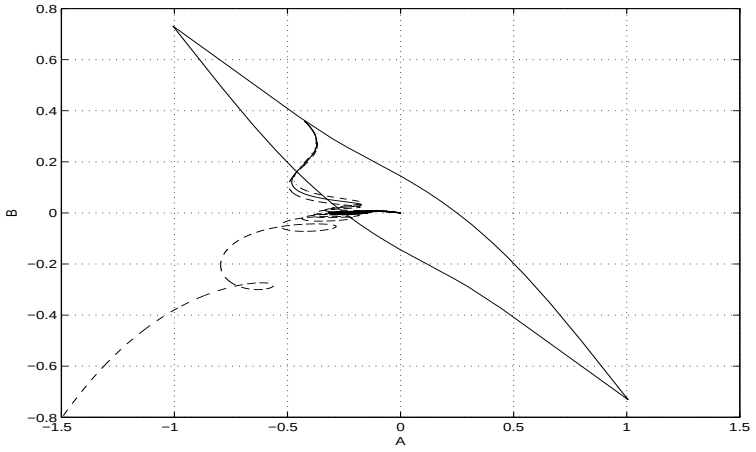


Figure 6: Even small amplitude noise can make big difference. We show the system with $\mu_1 = 0.7$, $\mu_0 = 0.1$, $k_1 = 1$, $l = 0$ and $\epsilon = 10^{-3}$. Three different trajectories are shown for $\omega = 2.4$ (dashed), 2.55 (solid) and 2.7 (dashed again). All have same initial perturbation. This illustrates the importance for small amplitude disturbances in the system.

6.2 Small amplitude noise system

In any physical system background noise of small amplitude is always present. In present case we have only considered marginally stable modes but the stable modes are still in the background and can be considered as background noise. Thus it is important to understand how system behaves in presence of small amplitude noise.

It would be very difficult to analyze the system with random noise. So we considered putting noise of particular frequency. This physically can be explained as the most dominant frequency in the noise. Ultimately we would like to consider the effect of random noise as well.

The modified equations with the noise for triple diffusive system can be written as,

$$\dddot{A} + \hat{\mu}_2\ddot{A} + \hat{\mu}_1\dot{A} + \hat{\mu}_0 A = k_1 A^3 - lA^5 + \epsilon A \sin(\omega t) \tag{46}$$

and for double diffusive system is,

$$\ddot{A} + \hat{\mu}_1\dot{A} + \hat{\mu}_0 A = k_1 A^3 - lA^5 + \epsilon A \sin(\omega t) \tag{47}$$

where, ϵ is atleast an order of magnitude smaller than anything else.

Figure 6 shows three different trajectories for $\omega = 2.4$, 2.55 and 2.7. A frequency of 2.4 makes the system unstable even though we started well within the basin of attraction of the unperturbed equations. For $\omega = 2.55$, we see the system go into an almost periodic orbit for a very long time before decaying back to origin. Figure 7 shows the behavior of A with time for this case. It is clear from this figure that we go into a orbit very close to periodic orbit for $\omega = 2.55$. The system decays down to origin for $\omega = 2.7$. It is clear the we can go to other states even after starting in the basin of attraction of origin with small noise in the system.

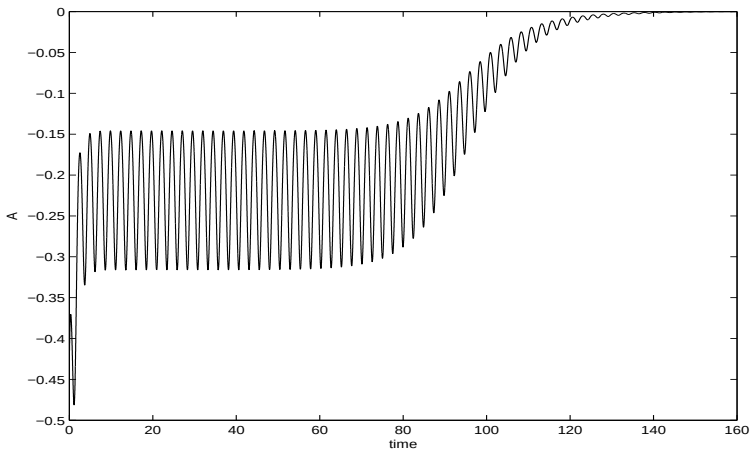


Figure 7: Nearly periodic orbit for $\omega = 2.55$ till $t \approx 80$ We can make system stay on this orbit longer by choosing a more suitable frequency.

It is also possible to go to infinity even if we start infinitesimally close to origin as shown in figure 8. There is range of ω values that can take the system out of the basin of attraction even with very small initial disturbances. This physically seems to suggest that background disturbances even of very small magnitude can make system unstable in linearly stable region.

We get a range of interesting behavior if we take $l \neq 0$ in eq 47. We started increasing ω from 0. For very small values, $\omega \leq 1.1$, almost no effect of the noise is seen on the system and for small initial disturbances the system finally goes to stable fixed point at origin. As ω is increased further the system goes away from origin into whats seems like a chaotic orbit as seen in fig 9.

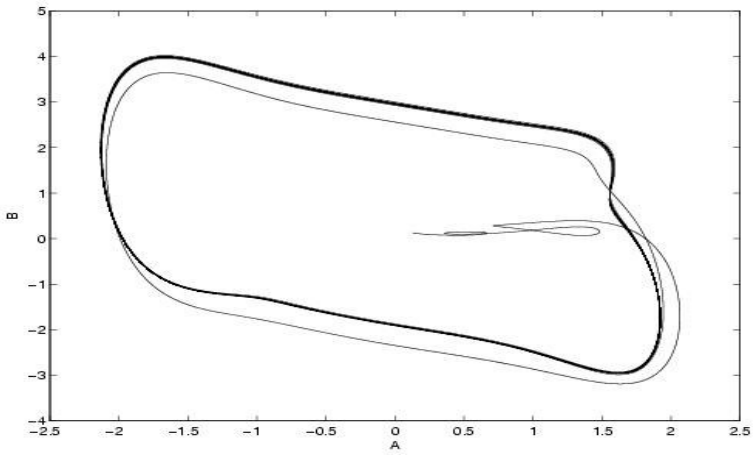


Figure 10: For $\omega = 1.9$ seems to go into a periodic orbit.

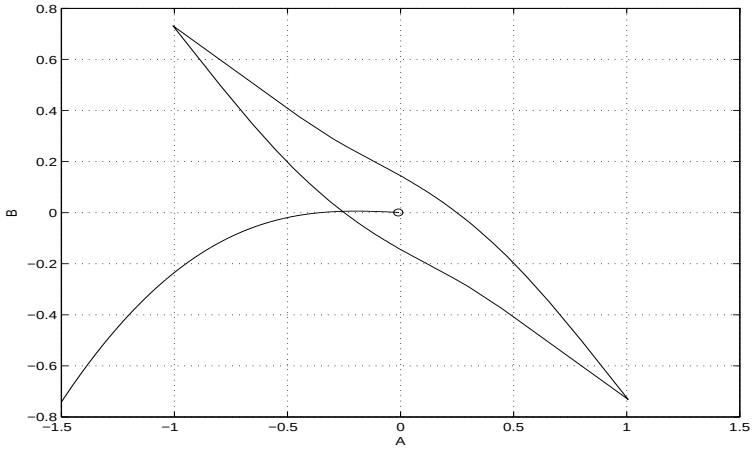


Figure 8: Initial disturbance was of the order of 10^{-3} and still we find a range of frequency of small amplitude noise that takes the system to infinity As we show here with $\mu_1 = 0.71$, $\mu_0 = 0.1$, $\epsilon = 10^{-3}$ and $\omega = 0.01$.

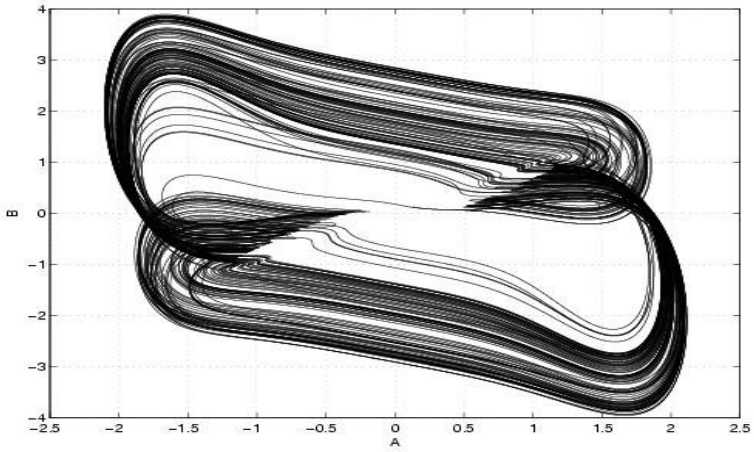


Figure 9: The system with $\mu_1 = 0.5$, $\mu_0 = 0.1$, $\epsilon = 10^{-3}$ and $l = 0.5$. The system above with $\omega = 1.2$ seems to go into a chaotic looking orbit.

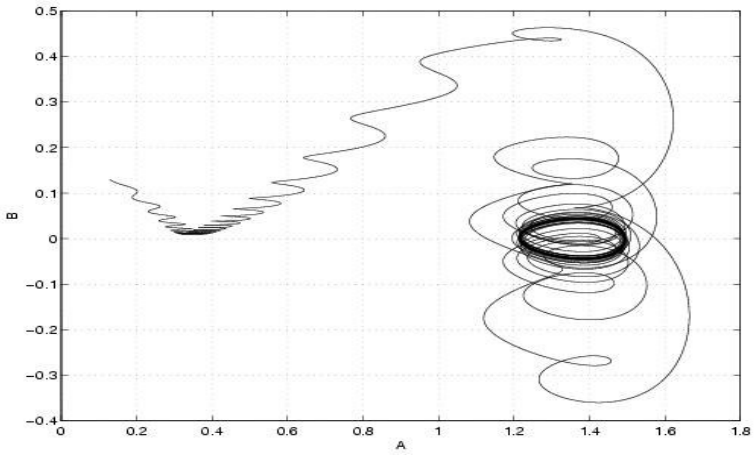


Figure 11: For $\omega = 10.0$ seems to go into a periodic orbit around the outer fixed point.

When the frequency is increased further we observe there is a small range of frequency that again take it to origin. Further increasing the frequency takes the system into a periodic orbit around the outer stable fixed point. This orbit is stable and system stays in this periodic orbit as shown in fig 10.

Increasing frequency further, we go in a regime where the system goes into periodic orbit around the outer fixed point. This periodic orbit is also stable, shown in figure 11. If the frequency is still increased further, the system goes to origin again. Further increase in frequency does not lead to any further change in this behavior.

7 Conclusion

Non-normal growth is observed for double and triple diffusion convection near the polycritical surface. This suggests that apart from shear flows non-normal growth can also be important in double diffusive convection as a pathway to turbulence, as suggested in Trefethen *et al.* (1993).

Non-normal growth along with the bootstrapping mechanism is not sufficient to take our system away from stable fixed points. We thus considered the possibility of noise along with the bootstrapping mechanism as considered in Gebhardt & Grossmann (1994). We have clearly demonstrated that small noise along with the non-normal nature of system leads to possible escape route out of the basin of attraction of stable fixed point for double diffusive systems.

We observe a range of behavior exhibited by the system as the frequency of the noise term is varied. This behavior appears similar to stochastic resonance in nonlinear systems (c.f. Locher *et al.* (1998)). We are able to observe range of frequencies that can take the system away from stable fixed points. We also observe that a very small disturbance with correct noise frequency can cause the system to go away from the basin of attraction. This supports the suggestions that rapid changes in ocean circulation may indeed be possible even with small disturbances to current environment (c.f. Wiesenfeld & Moss (1995) and

Ganopolski & Rahmstorf (2002)).

Results suggest that small noise have important effect even in linearly stable regime in non-normal systems as non-normality amplifies the effect of noise. Further work is need to consider effect of white noise and possible obtain a probability distribution of escape from basin for different noise levels.

8 Acknowledgments

I would like to thank Ed Spiegel for suggesting this problem, Norman Lebovitz for his valuable support throughout the summer, Joe Keller for much needed help in understanding all those papers, Louis Howard for his invaluable suggestions, George Veronis for introducing me to the pleasures of softball, Neil Balmforth and Stefan for keeping all of us on our toes. I would also like to thank Shilpa and Shreyas for all the wonderful meals and suggestions. Finally my fellow fellows for making this experience most wonderful.

References

Arneodo, A., Coullet, P. & Spiegel, E. 1985 The dynamics of triple convection. *Geophy. Astrophys. Fluid Dynamics* **31**, 1.

Baggett, J., Driscoll, T. & Trefethen, L. 1993 A mostly linear-model of transition to turbulence. *Phy. of Fluids* **7**, 833.

Chandrasekhar, S. 1961 Hydrodynamic and hydromagnetic stability. *Dover Publications Inc., Newyork* .

Coullet, P. & Spiegel, E. 1983 Amplitude equations for systems with competing instabilities. *SIAM J. Appl. Math* **43**, 776.

Ganopolski, A. & Rahmstorf, S. 2002 Abrupt glacial climate changes due to stochastic resonance. *Phys. Rev. Lett.* **88**.

Gebhardt, T. & Grossmann, S. 1994 Chaos transition despite linear-stability. *Phys. Rev. E* **50**, 3705.

Grossmann, S. 2000 The onset of shear flow turbulence. *Rev. Mod. Phys.* **72**, 603.

Locher, M., Cigna, D., Hunt, E., Johnson, G., Marchesoni, F., Gammaitoni, L., Inchiosa, M. & Bulsara, A. 1998 Stochastic resonance in coupled nonlinear dynamic elements. *Chaos* **8**, 604.

Shirtcliffe, T. 1967 Thermosolutal convection - observation of an overstable mode. *Nature* **213**, 489.

Stommel, H., Arons, A. & Blanchard, D. 1956 An oceanographical curiosity - the perpetual salt fountain. *Deep-Sea Research* **3**, 152.

Trefethen, L., Trefethen, A., Reddy, S. & Driscoll, T. 1993 Hydrodynamic stability without eigenvalues. *Science* **261**, 578.

Wiesenfeld, K. & Moss, F. 1995 Stochastic resonance and the benefits of noise - from ice ages to crayfish and squids. *Nature* **373**, 33.

Parametric instability of internal waves with rotation

Visweswaran Nageswaran
University of Massachussets,Amherst

1 Introduction

Internal gravity waves carry a lot of energy in the ocean and are known to be unstable. Their eventual breakdown thus has a significant effect in the oceans because the waves cause substantial mixing. Internal wave driven mixing not only affects the local circulation patterns but also energitically controls the global-scale meridional overturning circulation as suggested by previous results, and hence is vital for understanding of past, present and future climates([2],[4]). A substantial portion of internal wave energy is in the form of an internal tide. The internal tide is created by the interaction of the barotropic (astronomical) tide with rough topography. A variety of recent results have mapped out global locations of efficient internal tide generation. However, once generated, the internal tide does not necessarily break down and produce mixing near the generation site. Observations from the Hawaiian Ocean Mixing Experiment show that a low mode internal tide can propagate up to thousands of kilometers from it's generation site. The global pattern of tidal mixing is therefore strongly influenced by the dynamical instabilities that can efficiently drain energy from a propagating wave. Recent 3-D simulations of the mode1(M2) internal tide done by Mackinnon and Winters [5] suggests that the M2 tide loses energy at a particular latitude where the tidal frequency is twice the local inertial frequency. This raises a lot of interesting questions regarding the instabilities. For example, what is the growth rate of the instabilities at the crtical latitude?.What are the dynamic nature of those instabilities?.

Resonant wave triad interactions offer a framework to investigate the wave interactions causing the instability of the primary M2 tide. Neef and Mackinnon [3] have modelled the phenomenon as a system of one primary (tidal) wave and two close to subharmonic (M2/2) waves and show that as one approaches the critical latitude the rates of tansfer of energy get larger.However the above approach suffers from the drawback that at the critical latitude the frequency of the subharmonic waves appropaches the local inertial frequency. Therefore they cannot be treated as propagating waves anymore. In fact, the MacKinnon and Winters simulations show that the growing instabilities at this latitude look like pure inertial oscillations.

Motivated by their results, we have undertaken a series of analysis to look at the interaction between a large-scale propagating tide and small-scale near-inertial oscillations at a latitude where the tidal frequency is twice the local inertial frequency. our primary goal would be to study the growth rates of instabilities that drain energy away from the primary M2 internal tide at the critical latitude. In sections 2 and 3 we do a linear stability

analysis on the primary M2 internal tide within the framework of Boussinesq model with constant stratification frequency and study the growth rate of the instabilities at the critical latitude. In section 5 we propose an analytic approach that models the interaction of the primary internal tide with the instabilities and derive the growth rates for the fastest growing perturbations.

2 Numerical approach

We consider the inviscid Boussinesq equation with rotation and with constant stratification frequency N.

$$\frac{Du}{Dt} - fv = p_x \tag{1}$$
$$\frac{Dv}{Dt} + fu = -p_y \tag{2}$$
$$\frac{Dw}{Dt} - b = -p_z \tag{3}$$
$$u_x + v_y + w_z = 0 \tag{4}$$
$$\frac{Db}{Dt} + N^2 w = 0 \tag{5}$$

where (u, v, w) is the velocity,p the pressure , b the buoyancy and

$$\frac{D}{Dt} = \partial_t + u\partial_x + v\partial_y + w\partial_z$$

Linearising equations (1)-(4) we get

$$u_t - fv = -p_x \tag{6}$$
$$v_t + fu = -p_y \tag{7}$$
$$w_t + b = -p_z \tag{8}$$
$$b_t + N^2 w = 0 \tag{9}$$
$$\nabla.\mathbf{u} = \mathbf{0} \tag{10}$$

The M2 internal tide given by

$$U = a\cos(mx - \sigma t)\cos nz \tag{11}$$
$$V = \frac{af}{\sigma}\sin(mx - \sigma t)\cos nz \tag{12}$$
$$W = a\sin(mx - \sigma t)\sin nz \tag{13}$$
$$B = \frac{-N^2}{\sigma}\cos(mx - \sigma t)\sin nz \tag{14}$$

is a solution of (6)-(10). where σ is the M2 tidal frequency = 1/(12.4 hrs),$n = 2\pi$/ocean depth = $2\pi/4000$,m is a typical horizontal wavenumber = $2\pi/150km$ for midlatitudes and 'a' is a typical tidal velocity, on the order of 5 cm/s. Motivated by the tide propagating north from Hawaii, we're considering a tide propagating purely northward with infinite zonal wavenumber - hence there is no y dependence.

3 Floquet system governing the instability of the M1 internal tide

To investigate the stability of M2 internal tide we consider perturbations around this state.Let

$$u = U + u', v = V + v', w = W + w', b = B + b', p = P + p'. \tag{15}$$

Linearising the equations (1)-(5) for small perturbations u',v',w',b',p' gives (after dropping ' notation uniformly),

$$\begin{gathered}
\frac{Du}{Dt} + uU_x + wU_z - fv = -p_x \\
\frac{Dv}{Dt} + uV_x + wV_z + fu = -p_y \\
\frac{Dw}{Dt} + b = -p_z \\
\frac{Db}{Dt} + uB_x + wB_z + = 0 \\
u_x + v_y + w_z = 0
\end{gathered}$$

Where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + U\frac{\partial}{\partial x} + V\frac{\partial}{\partial y} + W\frac{\partial}{\partial z} \tag{16}$$

We now introduce the following scalings

$$\begin{aligned}
(x, y, z) &= (mx^*, my^*, nz^*) \\
(U, V) &= \frac{\sigma}{m}(U^*, V^*) \\
W &= \frac{\sigma}{n}W^* \\
B &= \frac{\sigma^2}{n}B^* \\
(u, v) &= \frac{\sigma}{m}(u^*, v^*) \\
w &= \frac{\sigma}{n}w^* \\
b &= \frac{N^2}{n}b^* \\
P &= \frac{\sigma^2}{m^2}p^*
\end{aligned}$$

The equations of perturbation become

$$\frac{Du}{Dt} - a\sin(x-t)\cos(z)u - a\cos(x-t)\sin zw - \frac{f}{\sigma}v = -p_x \tag{17}$$

$$\frac{Du}{Dt} + a\frac{f}{\sigma}\cos(x-t)\cos(z)u - a\frac{f}{\sigma}\sin(x-t)\sin(z)w + \frac{f}{\sigma}u = -p_y \tag{18}$$

$$\frac{Dw}{Dt} - a\cos(x-t)\sin(z)u - a\sin(x-t)\cos(z)w - \frac{N^2}{\sigma^2} = -\frac{n^2}{m^2}p_z \tag{19}$$

$$\frac{Db}{Dt} - a\sin(x-t)\sin(z)u + a\cos(x-t)\cos(z)w + w = 0 \tag{20}$$

$$u_x + v_y + w_z = 0 \tag{21}$$

$$\frac{D}{Dt} = \partial_t + a\cos(x-t)\cos z\partial_x + a\frac{f}{\sigma}\sin(x-t)\cos z\partial_y + a\sin(x-t)\sin z\partial_z$$

We now try a solution for (17)-(21) of the form

$$\begin{aligned} u &= u(t)\cos\gamma z e^{\imath(\alpha x+\beta y)} \\ v &= v(t)\cos\gamma z e^{\imath(\alpha x+\beta y)} \\ w &= w(t)\sin\gamma z e^{\imath(\alpha x+\beta y)} \\ b &= b(t)\cos\gamma z e^{\imath(\alpha x+\beta y)} \\ p &= p(t)\sin\gamma z e^{\imath(\alpha x+\beta y)} \end{aligned}$$

where the z dependence is chosen to meet the rigid lid top and boundary conditions.After substituting for u,v,w,b,p and taking

$$\partial_x(17) + \partial_y(18) + \partial_z(19)$$

we solve for p in terms of u,v,w,b and back substitute in equations (17)-(20).This gives us a system ordinary differential equation

$$X_t = AX \tag{22}$$

$$\begin{bmatrix} L(\alpha\beta f + a(\alpha^2 + \imath(\imath+\beta f)(\beta^2+\gamma^2 k))\cos z \\ \sin(t-x) - a\beta^2\gamma\sin(t-x)\sin z\tan\gamma z - \\ a\gamma^3 k\sin(t-x)\sin z\tan\gamma z + a\alpha\cos(t-x) \\ ((2\beta f - \imath\beta^2 + \gamma^2 k))\cos z + 2\imath\gamma\sin z\tan\gamma z)) \\ \\ -(L(\alpha^2 f + f\gamma^2 k + \imath a\alpha\beta(2\imath+\beta f)\cos z\sin(t-x) - \\ a\alpha\beta\gamma\sin(t-x)\sin z\tan\gamma z + a\cos(t-x) \\ ((-\imath\alpha^2\beta + f(\alpha^2-\beta^2+\gamma^2 k))\cos z - 2\imath\beta\gamma\sin z\tan\gamma z))) \\ \\ -(L\csc\gamma z(-\imath\gamma k\sin(\gamma z)(\beta f + a\alpha(2-\imath\beta f)\cos z\sin(t-x) + \\ a\alpha\gamma\sin(t-x)\sin z\tan\gamma z) + a\cos(t-x)(\alpha^2+\beta^2+\gamma^2 k)\cos(\gamma z) \\ \sin z + \gamma k\sin(\gamma z)(\alpha^2 - 2\imath\beta f)\cos z + 2\gamma\sin z\tan\gamma z)))) \\ \\ a\cot(\gamma z)\sin(t-x)\sin z \end{bmatrix}$$

$$\begin{bmatrix} L(\beta^2 f + f\gamma^2 k + \imath a\alpha^2\beta\cos(t-x)\cos z - \imath a\alpha\beta^2 f\cos z\sin(t-x) + a\alpha\beta\gamma\sin(t-x)\sin z\tan\gamma z) \\ \imath L(-(a\alpha(\alpha^2+\gamma^2 k)\cos(t-x)\cos z) + a\beta f(\alpha^2+\gamma^2 k)\cos z\sin(t-x) + \\ \imath(\alpha\beta f + a\gamma(\alpha^2+\gamma^2 k)\sin(t-x)\sin z\tan\gamma z)) \\ \\ \gamma k L(-\imath\alpha f - a\alpha\beta\cos(t-x)\cos z + a\beta^2 f\cos z\sin(t-x) + \imath a\beta\gamma\sin(t-x)\sin z\tan\gamma z) \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} aL(\alpha\sin(t-x)(-((-2\imath+\beta f)\gamma\cos z)+(2\beta f-\imath\gamma^2)\sin z\tan\gamma z)+ \\ \cos(t-x)(\alpha^2\gamma\cos z+(-\alpha^2+\beta^2+\gamma^2 k)\sin z\tan\gamma z)) \\ -(aL(\alpha\beta\cos(t-x)(-(\gamma\cos z)+2\sin z\tan\gamma z)+\sin(t-x)(\beta(-2\imath+\beta f)\gamma\cos z+ \\ (\imath\beta\gamma^2+f(\alpha^2-\beta^2+\gamma^2 k))\sin z\tan\gamma z))) \\ -(aL\csc\gamma z(-(\gamma(\alpha^2+\beta^2+\gamma^2 k)cos(\gamma z)\sin(t-x)\sin z)+ \\ \imath\sin\gamma z(\alpha\cos(t-x)((\alpha^2+\beta^2)\cos z+2\gamma k\sin z\tan\gamma z)- \\ \sin(t-x)((-\imath\beta^2+\beta^3 f+\alpha^2(-\imath+\beta f)+\imath\gamma^2 k)\cos z+\gamma(2\beta f-\imath\gamma^2)k\sin z\tan\gamma z)))) \\ -1+a\cos(t-x)\cos z \end{bmatrix}$$

$$\begin{bmatrix} \frac{\imath\alpha F^2\gamma L}{\sigma^2} \\ \frac{\imath\beta F^2\gamma L}{\sigma^2} \\ \frac{(\alpha^2+\beta^2)F^2 L}{\sigma^2} \\ a(-\imath\alpha\cos(t-x)\cos z+ \\ \sin(t-x)(\imath\beta f\cos z+ \\ \gamma\cot(\gamma z)\sin z)) \end{bmatrix}$$

and
$X = [u, v, w, b]^t$.

Since the matrix A is periodic with period 2π we can use floquet analysis to determine the stability/instability of the perturbations. The Floquet multipliers of the system then correspond to the growth rate of the perturbations. For a given x,y and z and for a given value of α, β, γ we can find the floquet multipliers which in turn determine the growth rate of the instabilities. From Floquet theory we know that there exists a constant matrix E which satisfies

$$X(\tau + 2\pi) = X(\tau)E, \forall\tau \tag{23}$$

The floquet multipliers are now defined as

$$\mu = \log\lambda$$

where λ is the eigen value of E.The real part of μ corresponds to the growth rate.If the absolute value of an eigen value is greater than unity then the perturbation is unstable. To determine E numerically let W(0) be the 4x4 identity matrix whose columns are the initial conditions for the system. Integrating (22) numerically from 0 to 2π let us denote by $W(2\pi)$ the matrix whose colums represent the solution of the system at time $t = 2\pi$. From (23) we get $E = W(0)^{-1}W(2\pi)$.ie,

$$E = W(2\pi)$$

4 Numerical results and discussion

Since we are interested in what the fastest growing modes are at the crtitical latitude we analyse the growth rates of the perturbations at the critical latitude. This corresponds to

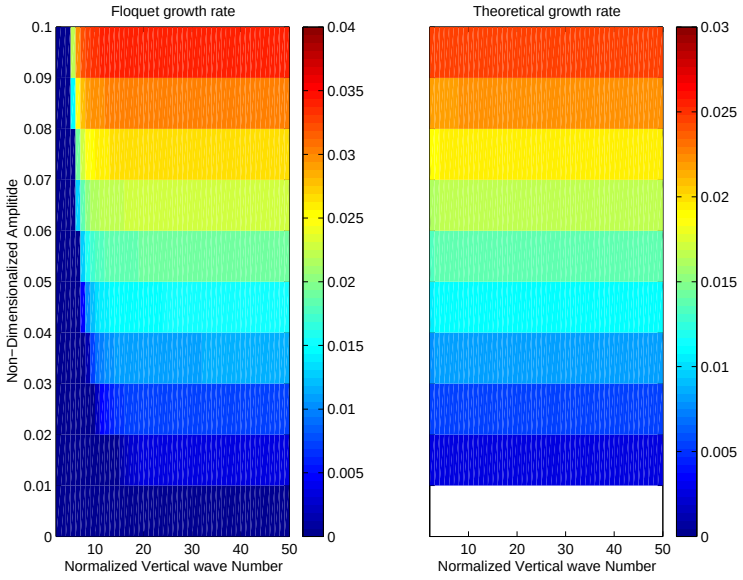


Figure 1: Theoretical growth rate vs Numerical growth rate

the case when $\frac{f}{\sigma} = 0.5$. We expect the growth rates to vary with depth. Since the waves and perturbations are uniform in the x,y direction we can fix $x = 0 = y$ and analyse the growth rate for different values of z. For the numerical calculation we fixed $\alpha = 1, \beta = 0$ and computed the growth rates numerically.The results are shown in figure 1 for z=0.02 fixed. The growth rate increases with the increase in amplitude of base wave. Also for a fixed amplitude the growth rate increases asymptotically to a constant value with increasing verical wave number for a given amplitude. The highest growth rates corresponds to wave numbers greater than ten. Infact for a fixed amplitude the growth rate shows a sudden increase for wave numbers greater than ten which suggests that the fastest growing instabilities have vertical scales which are ten times smaller than that of the original base wave. This agrees with the numerical simulations done by Winters and Mackinnon as well. However detailed numerical calculations need to be done to verify the results for a range of parameters α, β, γ and compare it with the theoretical prediction.

5 Analytical study of growth rate of instability at the critical latitude

We now do an analytical study of the growth rate of perturbations at the critical latitude,near the inertial frequency. We now introduce the following scalings $(u, v) = Rf(u^*, v^*), b = N^2Hb^*, p = N^2H^2p^*, t^* = ft$ where $R = \frac{NH}{f}$ is the Rossby deformation radius and H the depth of the ocean. The scales for x, y are given by $(x^*, y^*) = (\frac{x}{R}, \frac{x}{R})$ respectively. Numerical results, and observations from the Hawaii Ocean Mixing Experiment (HOME) both suggest that the strongest instabilities occur near the boundaries, where the tidal velocities

are largest. Hence the scale for z is given by

$$z = \frac{Z}{\epsilon H} \tag{24}$$

where $\epsilon^2 = \frac{a\sigma}{mNH}$ is small. We also use the hydrostatic approximation $N^2 \gg \sigma^2$. After dropping the * notation for convenience the perturbation equations become,

$$\begin{aligned}
\frac{Du}{Dt} + auU_x + awU_z - v &= -p_x \\
\frac{Dv}{Dt} + auV_x + awV_z + u &= -p_y \\
-\frac{p_z}{\epsilon} + b &= 0 \\
\frac{Db}{Dt} + uB_x + wB_z + &= 0 \\
u_x + v_y + w_z &= 0
\end{aligned}$$

Since the vertical scale of the perturbations is small compared to the horizontal scale(as suggested by the scale for z in (24)) we let $w = \epsilon w, b = \epsilon b$.Then p scales with ϵ^2 and the equations can be written as

$$\begin{gathered}
\tfrac{Du}{Dt} + auU_x + a\epsilon wU_z - v = -\epsilon^2 p_x \\
\tfrac{Dv}{Dt} + auV_x + a\epsilon wV_z + u = -\epsilon^2 p_y \\
-p_z + b = 0 \\
u_x + v_y + w_z = 0 \\
\tfrac{Db}{Dt} + auB_x + awB_z + \epsilon w = 0
\end{gathered}$$

If we choose $a = \epsilon^2 a_2$ where a_2 is an O(1) constant,then the equations can be written as

$$\begin{gathered}
\tfrac{Du}{Dt} + \epsilon^2 a_2 uU_x + \epsilon^3 a_2 wU_z - v = -\epsilon^2 p_x \\
\tfrac{Dv}{Dt} + \epsilon^2 a_2 uV_x + \epsilon^3 awV_z + u = -\epsilon^2 p_y \\
-p_z + b = 0 \\
u_x + v_y + w_z = 0 \\
\tfrac{Db}{Dt} + \epsilon a_2 uB_x + \epsilon^2 a_2 wB_z + \epsilon w = 0
\end{gathered}$$

where,

$$\frac{D}{Dt} = \partial_t + \epsilon^2 a_2 U\partial_x + \epsilon^2 a_2 V\partial_y + \epsilon a_2 W\partial_z \tag{25}$$

We can further approximate $\cos \epsilon z \approx 1$, $\sin \epsilon z \approx \epsilon z$ in lieu of the fact that the interaction takes place near the boundaries .

Numerical experiments reveal that the evolution of the instabilities takes place over several cycles of the primary M2 tide. Hence we assume that the variables depend on both

the fast time scale t and the slow time scale t_2. The appropriate slow time scale for the problem is $t_2 = \epsilon^2 t$. Note that the advective derivative now becomes

$$\frac{D}{Dt} = \partial_t + \epsilon^2 a_2 \left[\partial_{t_2} U \partial_x + V \partial_y + W \partial_z\right]$$

We now inroduce the complex field ,

$$q = u + iv$$

and

$$\xi = x + iy$$

The solution to the 0'th order equations can now be written from Ben-Jelloul,Young(97) [1] as follows

$$\begin{aligned} q_0 &= A_{zz} e^{-\imath t} \\ w_0 &= -M_\xi e^{-\imath t} + c.c \\ b_0 &= \imath M_\xi e^{-\imath t} + c.c \\ p_0 &= \imath A_\xi e^{-\imath t} + c.c \end{aligned}$$

where

$$M = A_z \tag{26}$$

$\mathrm{O}(\epsilon^2)$

$$q_{2,t_2} + iq_2 = -R_0$$

R_0 contains resonant terms i.e, terms proportional to $e^{-\imath t}$ and non-resonant terms. In order to be consistent with our assumptions we require that there are no resonantly growing terms which would cause the perturbation q_2 to grow larger than q_0. Hence equating the coefficient of R_0 that is proportional to $e^{-\imath t}$ to zero gives us an slow time evolution equation for A which is,

$$A_{zz,t_2} + \frac{\imath}{2}\nabla^2 A + \imath\gamma e^{imx} A^*_{zz} = 0 \tag{27}$$

where

$$\gamma = a_2 \frac{(1+\hat{v})m}{4}$$

We try a solution A of the form,

$$A = \sin(kz)\tilde{A}(x, t_2).$$

then A_0 and A_1 satisfies,

$$\tilde{A}(x, t_2) = A_0(t) + A_1(t) e^{imx}$$

and A_0, A_1 satisfies,

$$\begin{aligned} \dot{A}_0 + i\gamma \dot{A}^*_1 &= 0 \\ \ddot{A}_1 + i\left(\frac{m^2}{2k^2}\right)\dot{A}_1 - \gamma^2 A_1 &= 0 \end{aligned} \tag{28}$$

The solution of (28) is a linear combination of $e^{s_1 t}$ and $e^{s_2 t}$ where s_1 and s_2 are given by

$$s_1, s_2 = \frac{-i\frac{m^2}{2k^2} \pm \sqrt{4\gamma^2 - (\frac{m^2}{2k^2})^2}}{2}$$

The maximum of the real part of s_1 and s_2 determine the growth rate of the perturbations in slow time scale.

The predicted theoretical growth results suggests that the growth rate depends only on the vertical wave number k and the amplitude a_2 of the perturbation. Also the form of the solution given by (28) suggests to us that the fastest going perturbations have the same horizontal structure as that of the base wave. The growth rate increases with the increase of the amplitude and the vertical wave number. For a fixed amplitude the growth rate (see fig 1)shows a sudden increase for wave numbers greater than ten which illustrates that the vertical structure of fastest growing perturbationsare almost a factor of ten times smaller as compared to the base wave.

6 Conclusion

We have developed an anlaytic and a numerical model to study the growth rate of instabilities at the critical latitude.Both the thoeretical and numerical model predicts that the time period for the instabilities is roughly around 10-12 days which is relatively short. This has resulted in a significant improvement in our understanding of the role that internal tides plays in mixing. Most previous results(for example [2]) examine the role of internal tides for deep water mixing in the ocean. However, our new results and also other recent numerical results [5], now suggest that strong instabilities of the internal tide near the surface may lead to elevated tidal mixing in the upper ocean. Mixing in the upper ocean affects the rate heat, gasses and nutrients are moved around and stored on really short timescales (months to years), and so understanding upper ocean mixing is more relevant to ecosystem and climate evolution on short time scales whereas deep water mixing is generally important on longer time scales. Another fact that also emerges from our studies is that this is a very rich problem and offers a lot of scope for further work. An useful addition would be to carry out the anlaysis when the base wave is a plane wave which are exact solutions to (1)-(5). Also considering Beta plane effect and viscous effects will be an useful advance as well.

7 Acknowledgements

I would like to thank Jennifer Mackinnon for her patient guidance and inspiration throught the course of the summer project. I would also like to thank W.R.Young and Stefan Llewllyn Smith for their wonderful suggestions. I would also like to thank the other GFD fellows ,Shilpa Ghadge,Shreyas Mandre for their wondefrul friendship and last but not the least everyone else in Walsh cottage.

References

[1] M. Ben Jelloul and W. R. Young. "Propogation of near-inertial oscillations through a geostrophic flow." *J. Marine. Research* 55: 735-766 1997

[2] L. St. Laurent and C. Garrett "The role of internal tides in mixing the deep ocean." *J. Physical. Oceanography* 32: 2882-2899 2002

[3] Lisa Neef and J. A. Mackinnon "Triad Resonance as mehanism for Internal wave dissipation." *GFD 2004 Report*

[4] W. Munk and C. Wunsch "Abyssal recipes II: Energetics of tidal and wind mixing." *Deep sea research* 45 : 1977-2010 1998

[5] J. A. Mackinnon and K. B. Winters Private communication with J. A. Mackinnon

Triad Resonance as a Mechanism for Internal Wave Dissipation

Lisa Neef
University of Toronto

1 Introduction

Breaking of internal waves, i.e. the nonlinear transfer of energy from internal waves to turbulence, drives deep ocean mixing. Both the sources and sinks of internal wave energy in the world ocean, however, are spatially variable. It is estimated that roughly half of internal wave energy in the ocean is produced by the movement of the barotropic tide around topography, which generates internal waves with the same frequency. Where and when mixing due to the internal tide occurs, depends on two factors: the strength of the internal tide generated around a particular piece of topography, and the mechanisms that cause this internal tide to cascade into smaller scales, which then break and cause mixing. The magnitude and distribution in time and space of neither of these is well understood.

Much research is currently focused on how efficiently a particular topographic structure generates an internal tide. Furthermore, it is not well understood how far, in each case, a particular internal tide will propagate, and where and by what mechanisms it will deposit its energy. Recent studies(e.g. [1], [2]) have suggested that mixing due to internal wave breaking happens to a large extent in "hotspots" near rough topography.

MacKinnon and Winters [3] showed, using an idealized model, that an internal tide of M_2 frequency seems to efficiently break into waves of roughly half the M_2 frequency, and smaller vertical scale, suggesting that Parametric Subharmonic Instability (or PSI, i.e. the class of resonant wave-wave interactions wherein energy is transferred from a primary wave to two recipient waves of half the primary frequency and smaller vertical scale, and the range of interactions where the secondary waves are near this half-frequency) could be a significant mechanism of energy transfer to smaller scales. The efficiency of this interaction increased dramatically as the waves approached the critical latitude where the $M_2/2$ frequency equals the inertial frequency (f). The results of MacKinnon and Winters suggest not only that PSI could be a significant mechanism of energy transfer out of the internal M_2 tide, but also that this mechanism is strongly latitude-dependent. Though PSI has previously been thought too slow ([4]) to be an important mechanism for the transfer of energy into smaller-scales, MacKinnon and Winters argue that it might be much faster than previously thought in the nearfield of the generation site, where the internal tide is coherent.

In regards to latitude dependence, it is not clear why the efficiency of these interactions should increase so much near the critical latitude. One factor could be the fact that, near

the critical latitude, the waves with frequencies close to the subharmonic frequency have group velocities near zero, and, by breaking close to where they are generated, could create a "hotspot", where mixing is several times stronger than at lower latitudes. It is not yet known, however, how the efficiency of PSI interactions themselves actually changes with latitude, regardless of the group velocity of the resulting subharmonic waves. If PSI is indeed a significant factor in internal wave breakingg, its latitude dependence must be carefully examined.

The purpose of this paper is to find the intrinsic dependence of PSI interactions on latitude, or, more precisely, on the inertial frequency, f.

This will be done by rederiving the well-known triad equations for the case of internal waves with rotation of the earth (Section 2), and then integrating them numerically (Section 3). Section 4 then compares the results of the numerical integration to the established analytical solutions to the problem.

The greater purpose of this study is to help improve understanding of the spatial and temporal structure of tidal mixing, that is, how and where energy is transferred out of the internal tide. This is necessary in order to better understand the structure of the global ocean circulation, which requires both a map of the tidal generation sites, as well as a dynamical understanding of how much energy from the internal tide is lost to dissipation. Before PSI is dismissed as a dissipation mechanism, it must be considered for a coherent primary wave, and its latitude dependence examined. Understanding of this dependence is necessary for the correct modeling of energy dissipation of the internal tide.

2 Resonant Triad Model of Weakly Nonlinear Interactions

2.1 Derivation

We begin with the governing equations,

$$\frac{\partial u}{\partial t} + \mathbf{u}\cdot\nabla u - fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} \tag{1}$$

$$\frac{\partial v}{\partial t} + \mathbf{u}\cdot\nabla v + fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} \tag{2}$$

$$\frac{\partial w}{\partial t} + \mathbf{u}\cdot\nabla w = -\frac{1}{\rho_0}\frac{\partial p}{\partial z} - \rho' g \tag{3}$$

$$\frac{\partial \rho'}{\partial t} + w\frac{d\rho_0}{dz} = 0 \tag{4}$$

$$\nabla\cdot\mathbf{u} = 0. \tag{5}$$

One possible way to approximate weakly nonlinear interactions is via a resonant triad model. It is helpful to nondimensionalize by characteristic length, velocity, and time scales, so that different scales of motion can be examined. We thus scale the variables by typical velocity u_*, length scale k_*, and the time scale ω_*^{-1} of our primary wave, a typical propagating

internal tide. Thus,

$$(u, v, w)^{\mathrm{T}} = u_*(\tilde{u}, \tilde{v}, \tilde{w})^{\mathrm{T}} \tag{6}$$

$$x = \frac{\tilde{x}}{k_*} \tag{7}$$

$$t = \frac{\tilde{t}}{\omega_*} \tag{8}$$

where tildes indicate nondimensional quantities. The pressure-gradient terms on the right hand side must have the same dimensions as the left hand side. If we assume that the linear terms are dominant, the pressure-gradient terms must scale like the linear terms on the left hand side:

$$p = \frac{\rho_0 u_*}{k_* t_*}\tilde{p}. \tag{9}$$

We also define a nondimensional inertial and buoyancy frequencies, respectively,

$$f = \tilde{f}\omega_*, \tag{10}$$

$$N = \tilde{N}\omega_* \tag{11}$$

which are both scaled with respect to the linear wave motion.

Plugging (6) - (9) into (1) - (5), we have

$$\frac{d\tilde{u}}{d\tilde{t}} = -\frac{1}{\rho_0}\frac{d\tilde{p}}{d\tilde{x}} - \frac{u_* k_*}{\omega_*}\tilde{\mathbf{u}} \cdot \nabla\tilde{u} + \tilde{f}\tilde{v}, \tag{12}$$

and similar equations for $\tilde{v}$ and $\tilde{w}$. The nonlinear terms in these equations will be small if the quantity

$$\epsilon := \frac{u_*}{\omega_*/k_*} \tag{13}$$

is small. (13) shows that ϵ is the ratio of typical particle speed to the typical phase speed of the waves.

Equations (1)-(5) can be combined into a single equation:

$$\begin{aligned} &\partial_{tt}(-\nabla^2 - \partial_{zz})w - N^2\nabla^2 w - f^2\partial_{zz}w \\ -\epsilon[&\nabla^2(\mathbf{u}\cdot\nabla b) - \partial_t\nabla^2(\mathbf{u}\cdot\nabla w) + \partial_{xzt}(\mathbf{u}\cdot\nabla u) + \partial_{yzt}(\mathbf{u}\cdot\nabla v) - \\ &f\partial_{yz}(\mathbf{u}\cdot\nabla u) + f\partial_{xz}(\mathbf{u}\cdot\nabla u)] = 0 \end{aligned} \tag{14}$$

A useful tool for describing the impact of weakly nonlinear interactions on the linear dynamics is to define multiple time scales of motion. Having already defined a typical time scale of the primary wave, and nondimensionalized the governing equations by this time scale, we now introduce a comparatively *slow* timescale,

$$\tau = \epsilon t_* \tilde{t}, \tag{15}$$

with corresponding nondimensional time

$$\tilde{\tau} = \epsilon\tilde{t}. \tag{16}$$

$\tilde{\tau}$ is defined with respect to the magnitude of the nonlinear terms.

The time derivatives in (14) then become

$$\frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} + \epsilon \frac{\partial}{\partial \tau} \tag{17}$$

$$\frac{\partial^2}{\partial t^2} \rightarrow \frac{\partial^2}{\partial t^2} + 2\epsilon \frac{\partial^2}{\partial t \partial \tau} + \epsilon^2 \frac{\partial^2}{\partial \tau^2} \tag{18}$$

Substituting the right hand sides of (17) and (18) into (14), the $\mathcal{O}(\epsilon^0)$ balance is

$$\partial_{tt}(-\nabla^2 - \partial_{zz})w - N^2 \nabla^2 w - f^2 \partial_{zz} w = 0. \tag{19}$$

This is the linear component, and admits wave solutions like

$$w = w_i e^{i(k_i x + l_i y + m_i z - \omega_i t)} = \Sigma w_i e^{i\theta_i}, \tag{20}$$

with ω_i subject to the dispersion relation

$$\omega_i^2 = \frac{N^2 \ \ k_i^2 + l_i^2 \ \ + f^2 m_i^2}{k_i^2 + l_i^2 + m_i^2} \tag{21}$$

Substituting the linear solution (20) into the linear parts of (1) - (5), we can find polarization relations which relate u, v, p, and b to the prognostic variable, w:

$$u = g_u(w) = \frac{m(ik\omega - fl)}{i\omega(k^2 - l^2)} \ \ w \tag{22}$$

$$v = g_v(w) = \frac{m(ik\omega - fk)}{i\omega(k^2 - l^2)} \ \ w \tag{23}$$

$$p = g_p(w) = \frac{\omega^2 - N^2}{m\omega} \ \ w \tag{24}$$

$$b = g_b(w) = \frac{-iN^2}{\omega} w \tag{25}$$

The $\mathcal{O}(\epsilon)$ balance is

$$2\frac{\partial^2}{\partial t \partial \tau}(\nabla^2 + \partial_{zz})w = -\nabla^2(\mathbf{u} \cdot \nabla b) - \partial_t \nabla^2(\mathbf{u} \cdot \nabla w) + \partial_{zxt}(\mathbf{u} \cdot \nabla u) + \partial_{zyt}(\mathbf{u} \cdot \nabla v) - f\partial_{zy}(\mathbf{u} \cdot \nabla u) + f\partial_{zx}(\mathbf{u} \cdot \nabla v), \tag{26}$$

where ∇^2 is the horizontal Laplacian.

In order to model the nonlinear interaction as described by (26), we will seek solutions like

$$w = \sum z(\tau) e^{i(k_i x + l_i y + m_i z - \omega_i t)} + \sum z_i^*(\tau) e^{-i(k_i x + l_i y + m_i z - \omega_i t)} = \sum z_i(\tau)_i e^{i\theta_i} + c.c. \tag{27}$$

(27) is a sum of solutions to the linear problem. The complex amplitudes $z_i(\tau) = w_i(\tau) e^{i\phi_i}$ are allowed to vary on the slow timescale, to reflect the nonlinear interactions between the component waves.

(27) can then be substituted into (26) . Individual terms on the right hand side will then look like

$$\begin{aligned}
\mathbf{u}\cdot\nabla s &= u_1(\frac{\partial s_2}{\partial x}+\frac{\partial s_3}{\partial x}+...)+u_2(\frac{\partial s_1}{\partial x}+\frac{\partial s_3}{\partial x}+...)+v_1(\frac{\partial s_2}{\partial y}+\frac{\partial s_3}{\partial y}+...)+... \\
&= u_1\left(k_2 s_2+k_3 s_3+...\right)+u_2\left(k_1 s_1+k_3 s_3+...\right)+v_1\left(l_2 s_2+l_3 s_3+...\right)+... \\
&= \sum_i u_i \sum_{j\neq i} k_j s_j+\sum_i v_i \sum_{j\neq i} l_j s_j+\sum_i w_i \sum_{j\neq i} m_j s_j \qquad (28)
\end{aligned}$$

where s is can be one of the four system variables u, v, w, b.

Using the polarization relations (22)-(25), individual terms in (28) can be written in terms of the amplitudes z_i and phases θ_i of each component wave, e.g.

$$u_i k_j s_j = z_i g_u\left(z_i\right) e^{i\theta_i} k_j g_s\left(z_j\right) z_j e^{i\theta_j}. \qquad (29)$$

Consequently, derivatives of terms like $\mathbf{u}\cdot\nabla s$ will bring down factors of k_i+k_j, l_i+l_j, m_i+m_j, and $\omega_i+\omega_j$.

Plugging (27) into the left hand side of (26) gives

$$2i\sum_k \omega_k \;\; k_k^2+l_k^2+m_k^2 \;\; \frac{\partial z_k}{\partial \tau} e^{i\theta_k}. \qquad (30)$$

If we now consider the evolution of a particular wavenumber component (k_0, m_0, ω_0) by averaging both the left hand side and the right hand side over the period of that wave, the terms on the right hand side that will balance the left hand side will be ones where

$$\theta_i+\theta_j=\theta_0. \qquad (31)$$

The simplest possible nonlinear interaction we can consider is therefore a triad of waves where (31) holds, or, more specifically, where

$$\begin{aligned}
k_1+k_2 &= k_0 \qquad (32)\\
l_1+l_2 &= l_0 \qquad (33)\\
m_1+m_2 &= m_0 \qquad (34)\\
\omega_1+\omega_2 &= \omega_0. \qquad (35)
\end{aligned}$$

The equations for this interaction found by truncating (26) for the triad of waves, so that (27) becomes

$$w=\sum_{i=1}^{3} z_i(\tau)e^{i\theta_i(t)}+c.c.=\sum_{i=1}^{3} w_i(\tau)e^{i(\theta_i(t)+\phi_i(\tau))}+c.c. \qquad (36)$$

Plugging (36) into (26) and factoring out the complex exponentials in the individual terms as in (29) and (30), and applying the resonance condition (31), such that the phases of the individual waves, $e^{i\theta_i}$ divide out, we are left with equations for the evolution of the

complex wave amplitudes, $z_i(\tau)$, on the "slow" time scale (see Appendix for details). The resulting model is then given by

$$\dot{z_0} = -i\Gamma_0 z_1 z_2 \tag{37}$$

$$\dot{z_1} = -i\Gamma_1 z_0 z_2^* \tag{38}$$

$$\dot{z_2} = -i\Gamma_2 z_0 z_1^* \tag{39}$$

where the ∂_τ time derivative has been replaced with a dot. For each wave, the corresponding interaction coefficient Γ_i is a function of the wavenumbers of the two interacting waves. The interaction coefficients are derived in the appendix.

2.2 Comments

1. It is important to remember that the system (37) - (39) describes the coupled fluctuations of the amplitudes of the three waves, not the waves themselves. However, the individual fields of each variable can be found, for each wave, by substituting the amplitudes w_i of each wave, and the appropriate wave numbers, into the polarization relations (22) - (25). The frequency of the actual waves will be proportional to that of the triad interaction by a factor of ϵ.

2. The appropriate combinations of wavenumbers cannot be arbitrary, but must be chosen so that both the dispersion relation (21) and the resonance condition for each dimension (32) - (35) are satisfied. Thus, for example, given some horizontal structure described by (fixed) horizontal wave numbers k_i and l_i, decreasing ω_i means that the corresponding set of available vertical wavenumbers m_i will have larger values.

3. While the triad equations (37) - (39) can be easily integrated numerically, it is trickier to find resonant sets of wavenumbers and frequencies which also satisfy the dispersion relation. For example, suppose that the frequency, ω_0 and spatial wavenumbers k_0, l_0, m_0 of the primary wave are known. This leaves eight unknown parameters, but the system is then constrained by four resonance equations (32) - (35) and two dispersion relations, leaving two degrees of freedom. Thus, given the primary wave, it is only possible to choose two more wavenumbers in either one of the two secondary waves, say, for example, k_1 and l_1. In this case, k_2 and l_2 will be given by the appropriate resonance conditions, and m_1, m_2, ω_1, and ω_2 must be found by solving the system of equations given by the remaining resonance conditions (34), (35), and the dispersion relations (21) for wave 1 and wave 2. Such a system is impossible to solve analytically, but can be solved numerically.

4. It is easily seen from the triad equations (37) - (39) that the amplitude of energy exchange for each wave is proportional to the interaction coefficient for that wave – the wave with the smallest interaction coefficient exchanges the least amount of energy. It is also clear from the triad equations that the amplitude of energy exchange also depends on the initial amplitudes of the waves themselves, and their phases relative to each other.

5. The rate of change of each wave amplitude, and therefore the triad interaction frequency, is also a function of the initial amplitudes, and the interaction coefficients. It is this interaction frequency, and its dependence on the three sets of wavenumbers, on which this study focuses. It will be computed numerically in section 3, and analytically in section 4.

2.3 Methodology

In this study, we are interested in systems where the secondary waves (waves 1 and 2) have horizontal structures similar to that of the primary wave (wave 0), frequencies that are fractions of the primary wave frequency, and the appropriate vertical structures that result from satisfying both the dispersion relation (21) and the resonance conditions (32) - (35) . We therefore fix the parameters of the primary wave, with $\omega_0 = 1$ and k_0, l_0, and m_0 equal to three values given in Table 1. In order to simplify the problem, we choose $l_0 = l_1 = l_2 = 0$, thus reducing the problem to two dimensions while preserving the presence of the planet's rotation. This leaves one degree of freedom, which we choose to be ω_1. Now, varying ω_1, the corresponding ω_2, k_1, and k_2, m_1, and m_2 can be computed from the problem's constraints. We are thus in effect varying the frequency of the secondary waves and computing the corresponding horizontal and vertical structures.

For a given ω_1, there exists a set of possible combinations of m_1 and k_1, with consequent parameters for wave 2 given by the resonance condition. This is shown graphically in Figure 1, a plot of the vertical wavenumbers m_1 that satisfy each choice of ω_1. This so-called resonant trace has four different branches, two of which have large vertical wavenumbers as $\omega_{1,2} \to \omega_0$. The interaction coefficients in (37)-(39), and thus the interaction frequency, will be different for each branch.

The interaction frequency along these resonance traces will also depend on the other parameters in the dispersion relation, namely the buoyancy frequency, N, and the inertial frequency, f. Since the focus of this study is on the effect of latitude on the triad interaction, N will be kept constant ($N = 10$, in nondimensional units), while f will vary from 0 to $\omega_0/2$. For given $f > 0$, only those triads with $\omega_{1,2} \geq f$ will be allowed by the dispersion relation, thus truncating the resonant trace outside of the interval $[f, \omega_0 - f]$.

3 Numerical Results

Figure 2 shows the exchange of energy between three waves satisfying the resonance conditions, for a simple case where $f = 0$ and the primary wave is about an order of magnitude larger than the two daughter waves. The wavenumbers of the resonant triad in this example are given in Table 1. The corresponding fluctuation of a single wave, wave 0, is shown in Figure 3, where we have chosen $\epsilon = 0.05$. Figure 4 shows the composite wave field $w(x, z)$ for the same example, at $t = 0$ and again at a time halfway through the period of the wave interaction. At the half-period time, the vertical structure is considerably smaller, simply because Wave 2 ($m_2 = 13.36$) is dominant.

In this example, the frequency of energy exchange is about 0.75, on the nondimensional slow timescale. This means that exchange of energy between the triad members happens at a fraction of 0.75ϵ of the frequency of the primary wave, ω_0.

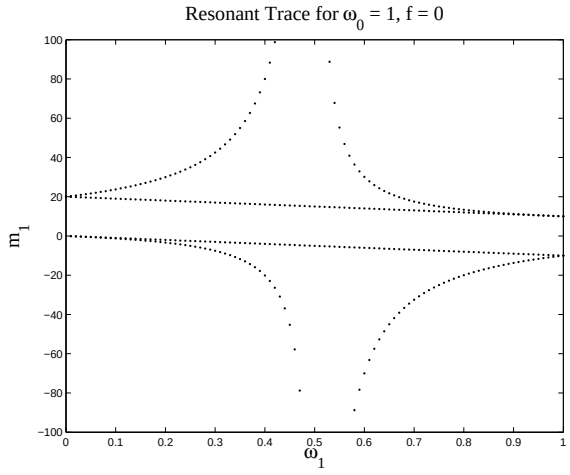


Figure 1: Pairs of frequency and vertical wavenumber satisying the resonance condition and dispersion relation, for $\omega_0 = 1$ and $l_0 = l_1 = l_2 = 0$.

Wave	Initial Amplitude w_i	Initial Phase ϕ_i	ω	k	m	Γ_i
0	1.0	3	-1	-1	10	-40.3
1	0.04	5	0.49	.12.1	-246	-4.74
2	0.02	5	0.51	-13.1	256	-5.34

Table 1: Initial values, wavenumbers, and interaction coefficients of the component waves in a sample resonant triad.

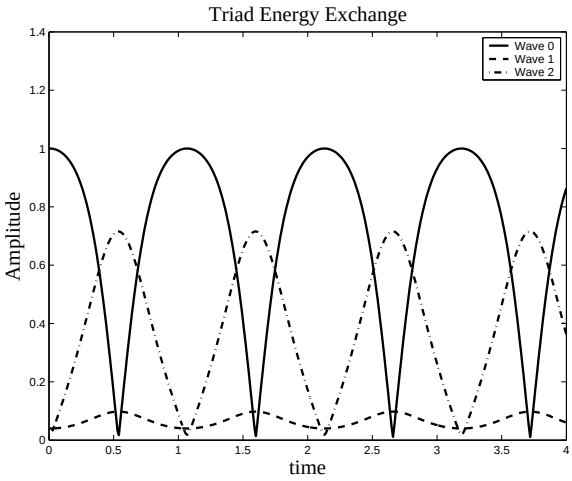


Figure 2: Exchange of energy between the three waves in a resonant triad. In this example, $f = 0$, and the wave frequencies are $\omega_0 = 1$, $\omega_1 = 0.2$, $\omega_2 = 0.8$. Initial wave amplitudes and phases are given in Table 1.

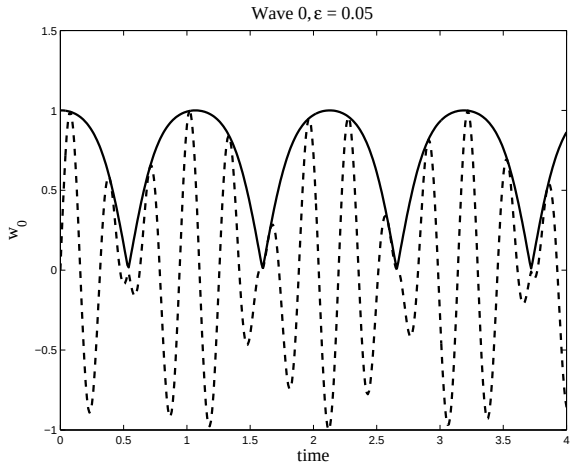


Figure 3: Amplitude fluctuation of the primary wave, for a triad oscillation with $f = 0$, and the wave frequencies and amplitudes as in Fig. 2.

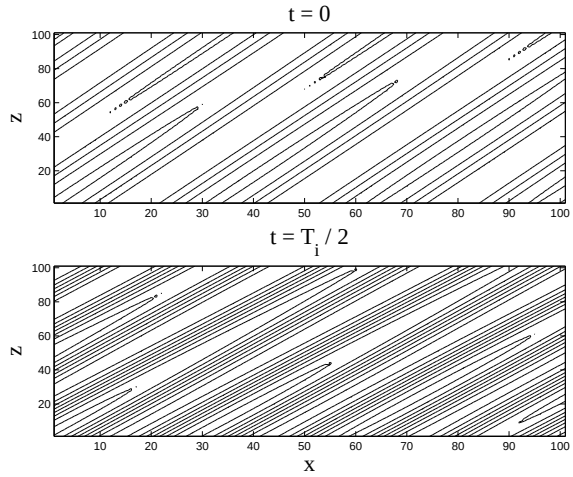


Figure 4: Fields of the variable w as a composite of the three interaction waves in the triad, at the initial time (left) and at $t = T/2$, where T is the period of the triad interaction.

Figure 5 examines the interaction frequency for triads with $\omega_0 = 1$, and values of ω_1 varying over the interval $[f, \omega_0/2]$ (with ω_2 changing in the other direction, to satisfy (35)). This plot has different characteristic curves, corresponding to the different branches of the resonant trace (Fig. 1). As shown in Figure 1, for a given value of ω_1, and taking $l_0 = l_1 = l_2 = 0$, there are typically three or four possible wave number combinations that satisfy both the dispersion relation and the resonance condition. Hence, a plot of the interaction frequency as a function of ω_1 has several different characteristic curves of interaction frequency. Four different increasing values of f are shown. Triads where $m_1 > 5m_0$ are indicated with an x. These indicate daughter waves with small vertical structure, the waves most likely to break and cause mixing.

For $f = 0$ the interaction frequency varies slowly with $\omega_{1,2}$, with a *minimum* at $\omega_{1,2} = \omega_0/2$. For increasing f, however, one of the resonant curves develops a sharp rise in the interaction frequency as $\omega_{1,2} \to \omega_0/2$. These are precisely those resonant triads where $m_1 > 5m_0$. As f approaches $\omega_0/2$ (the critical latitude), the band of available values of $\omega_{1,2}$ becomes smaller. This result is well known. It is surprising, however, that the overall interaction frequency seems to increase with increasing f.

Figure 6 shows the interaction frequency as a function of the three interaction coefficients, for the case $f = 0$. For each particular Γ_i, there exists usually more than one set of corresponding Γ_j, Γ_k, and thus more than one corresponding interaction frequency, making these figures difficult to interpret.. The dependence of the interaction coefficients on the resonant sets of wavenumbers (see Appendix), is similarly complicated. We now appeal to known analytical solutions to the triad equations to shed light on why the above numerical results are what they are.

4 Analytical Solution

It is possible to find an analytic solution to the system of triad equations (37) - (39), by deriving the system's invariants and transforming it into a single ODE.

First, it is helpful to transform the wave amplitudes so as to make the equations symmetric. This can be done by defining three "stretched" new wave amplitudes,

$$z_0 = \frac{\zeta_0}{2\sqrt{\Gamma_1\Gamma_2}} \tag{40}$$

$$z_1 = \frac{\zeta_1}{\sqrt{2\Gamma_0\Gamma_2}} \tag{41}$$

$$z_2 = \frac{\zeta_2}{\sqrt{2\Gamma_0\Gamma_1}} \tag{42}$$

The triad interaction equations then become

$$\dot{\zeta}_0 = -i\zeta_1\zeta_2 \tag{43}$$

$$\dot{\zeta}_1 = -\frac{i}{2}\zeta_0\zeta_2^* \tag{44}$$

$$\dot{\zeta}_2 = -\frac{i}{2}\zeta_0\zeta_1^*, \tag{45}$$

plus the three complex conjugate equations.

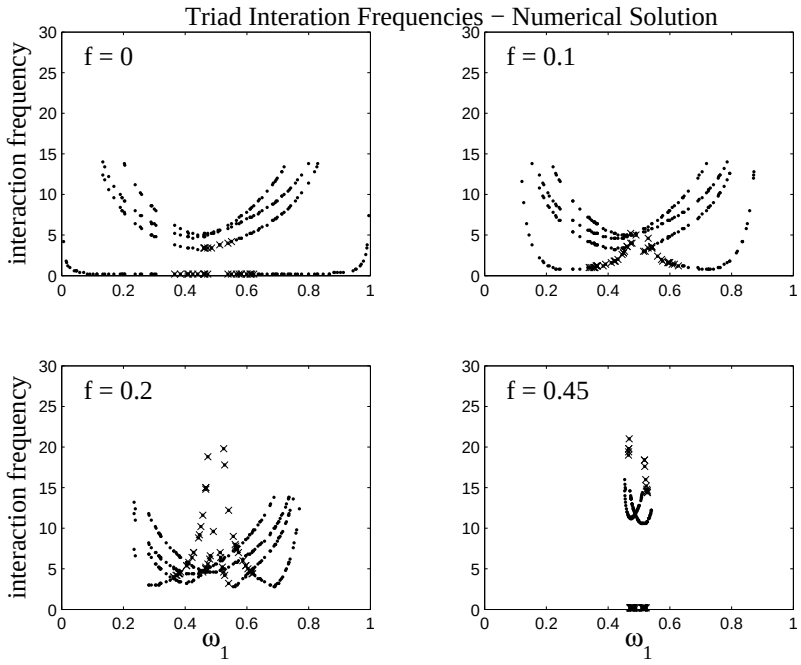


Figure 5: Interaction frequency for all possible resonant triads over a range of values of $\omega_{1,2}$, with everything else held constant. Points marked with an x denote triads where the vertical wavenumer of wave 1 is greater than five times that of wave 0.

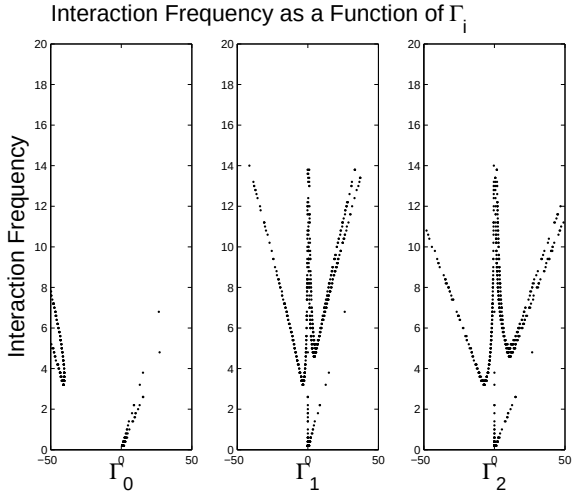


Figure 6: Interaction frequencies as a function of the three interaction coefficients, computed for a range of resonant triads with $f = 0$ and $\omega_0 = 1$.

Two invariants to the system are easily found. One is the energy,

$$|\zeta_0|^2 + |\zeta_1|^2 + |\zeta_2|^2 = A_0, \tag{46}$$

and another is given by one of the Manley-Rowe relations:

$$|\zeta_1|^2 - |\zeta_2|^2 = A_1. \tag{47}$$

Taking the time derivative of (43), and substituting (43) - (45) and their complex conjugates, and (46), gives

$$\ddot{\zeta_0} = -i\frac{d(\zeta_1\zeta_2)}{dt} = -\frac{1}{2}\zeta_0 \ \ A_0 - |\zeta_0|^2 \ \ . \tag{48}$$

This equation can be integrated twice to find an equation for $\zeta_0(t)$. (48) can be separated into its complex and imaginary components. Letting

$$\zeta_0 = u_0 e^{i\phi}, \tag{49}$$

(48) becomes

$$\ddot{u_0} + i\dot{u_o}\dot{\phi_0} + i\ddot{\phi_0}u_0 + i\dot{\phi_0}\dot{u_0} - \dot{\phi_0}^2 u_0 = -\frac{1}{2}u_0 \ \ A_0 - u_0^2 \ \ . \tag{50}$$

The imaginary component of (50) is easily integrated:

$$\begin{aligned} \frac{\ddot{\phi_0}}{\dot{\phi_0}} + 2\frac{\dot{u_0}}{u_0} &= 0 \\ \rightarrow \ln\dot{\phi_0} + \ln u_0^2 &= A_2 \\ \rightarrow \dot{\phi_0}u_0^2 &= A_3 \end{aligned} \tag{51}$$

The real component of (50) can now be rewritten using (51):

$$\ddot{u_0} - \frac{A_3}{u_0^3} = -\frac{1}{2} \ \ A_0 u_0 - u_0^3 \ \ . \tag{52}$$

This can be integrated as well, to give

$$\dot{u_0}^2 + \frac{A_3^2}{u_0^2} = -\frac{A_0}{2}u_0^2 + \frac{1}{4}u_0^4 + A_4. \tag{53}$$

Now writing $x = u_0^2$, we have

$$\begin{aligned} \dot{x}^2 &= x^3 - 2A_0x^2 + 4A_4x - 4A_3^2 \\ &=: -2\pi(x). \end{aligned} \tag{54}$$

Now (54) can be written as an integrable ODE, of the form

$$\frac{1}{2}\left(dx/dt\right)^2 + \pi(x) = 0,$$

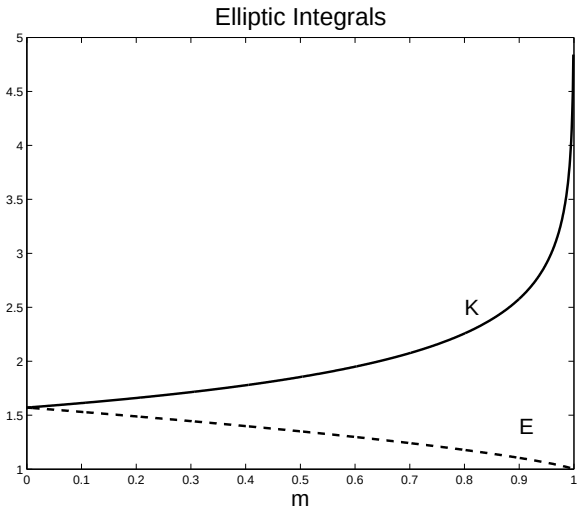


Figure 7: The complete elliptic integrals K and E as a function of the parameter m.

where $\pi(x)$ is a cubic function, with roots $a > b > c$.

If $\pi(x)$ has three real roots and $b < x(0) < a$, solutions $x(t)$ oscillate in a "potential well" bounded by the roots a and b. These oscillatory solutions are proportional to Jacobi Elliptic functions, with a period of

$$T = 2\sqrt{\frac{6}{a-c}}\left(cK(m) + (a-c)E(m)\right), \tag{55}$$

where K and E are the complete elliptic integrals of the first and second kind, respectively, and

$$m = \frac{a-b}{a-c}. \tag{56}$$

These are plotted in figure 7.

While this expression for the period is somewhat complicated, it can be interpreted in terms of the roots of $\pi(x)$, or the shape of the potential well. Assuming, for example, that the third root, c, stays constant, and letting $c = 0$, if the distance between the two largest roots increases (that is, if the size of the potential well increases), the parameter m will increase as well, E will go to zero, and T will decrease. If the distance between the largest and smallest root $(a - c)$ increases relative to $a - b$, then $m \to 0$, which means $K, E \to \pi/2$, and, again, T will decrease. The interaction frequency, therefore, can be interpreted in terms of the shape of $\pi(x)$.

The roots of $\pi(x)$ depend on the invariants in (54), which in turn depend on the initial wave amplitudes, as well as the interaction coefficients, which of course also depend on f. Figure 8 shows $\pi(x)$ plottted for $\omega_1 = .4$, and two values of f. As f approaches ω_1, the shape of $\pi(x)$ changes for one of the four solutions– the distance between the largest and smallest root becomes bigger, corresponding to a smaller interaction period (and higher interaction frequency).

Figure 9 shows the interaction frequency, computed analytically, for a range of ω_1, at four different values of f approaching the critical latitude, as in Figure 5.

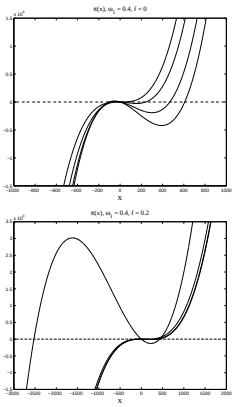


Figure 8: The potential function given in (54), for all 4 sets of resonant triads with $\omega_0 = 1$, and $\omega_1 = 0.2$. $f = 0$ in the top panel, and $f = 0.2$ in the bottom panel.

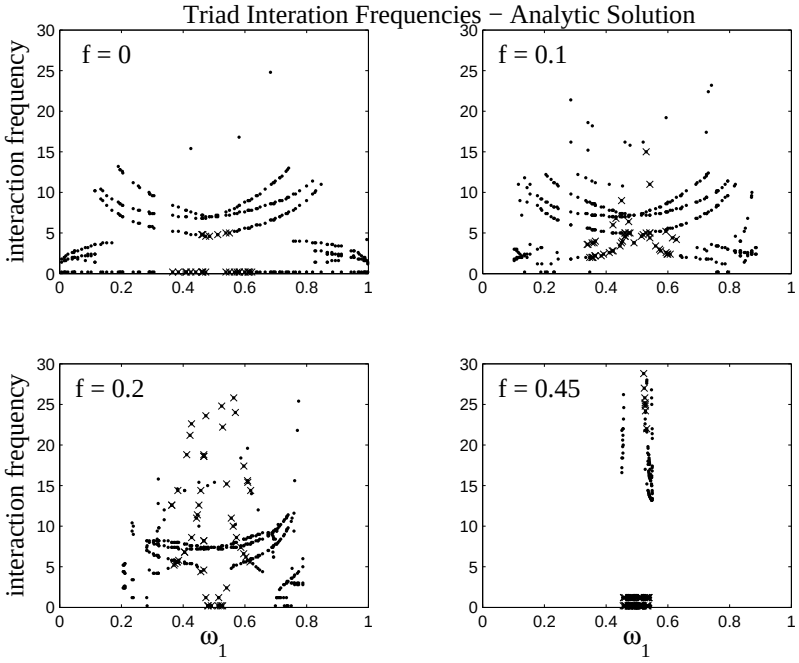


Figure 9: As in Fig. 5, but for the analytical solution given by (**??**): Interaction frequency for all possible resonant triads over a range of values of $\omega_{1,2}$, with everything else held constant. Points marked with an x denote triads where the vertical wavenumer of wave 1 is greater than five times that of wave 0.

As in the numerical integration of the triad equations and the above analytical reasonin, the interaction frequency at the subharmonic peak increases as f increases – thus showing that the behavior found in section 4 was not simply a numerical peculiarity.

5 Conclusions

This study has investigated the rate of energy transfer via triad resonance out of the internal M_2 tide .

- We have shown numerically that the efficiency of PSI increases as the waves approach the critical latitude where the Coriolis parameter f equals half the primary frequency, ω_0.
- This property is not simply a consequence of the fact that $\omega_0/2$ waves have zero group velocity at the critical latitude, but seems to be an intrinsic property of the triad equations.
- Since the triad equations are integrable, and analytical solution can be found, and the relationship between the system parameters, particularly f, and the interaction frequency can be examined more closely.

This study is a simple exercise intended to cast the more or less established and known properties of PSI into the context of the internal tide. Several things remain to be examined. A closer examination of the analytical solution and its relationship to f may be possible, especially if it is possible to cast the potential function $\pi(x)$ and its roots in a more transparent form. The problem should also be extended to the three dimensional ($l_i \neq 0$) case. The realistic numbers corresponding to the nondimensional results shown above (that is, an estimate of ϵ) also need to be considered.

6 Acknowldegements

I would like to thank the GFD coordinators and staff for giving me the opportunity to participate this summer, for teaching me a thousand things about tides, how to be a scientist, and softball. I would especially like to acknowledge Jen MacKinnon for all her help, especially the mega mathletics sessions at Walsh Cottage, late night pizza/chocolate provisions, endless patience and tons of enthusiasm. Thank you to Leo Maas for explaining Jacobi Elliptic functions to me, and for lots of helpful comments. Lots and lots of props also to Team Lazy (you know who you are) for consistently bringing the funny, even at 3 a.m. Thanks to all the amazing fellows, who's company I'm still intimidated to be in, and to Shreyas and Shilpa, who should have lost patience with me, but didn't. I'm knitting hats for everyone.

7 Appendix

To derive the triad equations (37)-(39), we substitute (36) into (26), and apply the resonance condition,

$$\theta_0 = \theta_1 + \theta_2. \tag{57}$$

The modulation of wave 0 by the other two waves can be found by plugging

$$z_0(\tau)e^{i\theta_0(t)}, \tag{58}$$

where $z_0(\tau)$ is the slowly-varying, complex amplitude of wave 0, into the left hand side of (26), and expanding each term on the right hand side in terms of waves 1 and 2. The left hand side then becomes

$$2\frac{\partial^2}{\partial t \partial \tau}(\nabla^2 + \partial_{zz})z_0 e^{i\theta_0} = 2i\omega_0 \;\; k_0^2 + l_0^2 + m_0^2 \;\; \frac{\partial z_0}{\partial \tau}e^{i\theta_0}. \tag{59}$$

On the right hand side of (26), the first gradient term becomes

$$\begin{aligned} -\nabla^2(\mathbf{u}\cdot\nabla\mathbf{b}) &= -\nabla^2 \;\; u_1\frac{\partial b_2}{\partial x} + v_1\frac{\partial b_2}{\partial y} + w_1\frac{\partial b_2}{\partial z} + u_2\frac{\partial b_1}{\partial x} + v_2\frac{\partial b_1}{\partial y} + w_2\frac{\partial b_1}{\partial z} \\ &= -\nabla^2 i\left[(u_1k_2 + v_1l_2 + w_1m_2)b_2 + (u_2k_1 + v_2l_1 + w_2m_1)b_1\right]. \end{aligned}$$

Substituting the polarizations relation between u and w, (22), and between b and w, (25), the right hand side becomes

$$-i\nabla^2 A_0 w_1 w_2,$$

where

$$\begin{aligned} A_0 w_1 w_2 &= [k_2 g_u(w_1) + l_2 g_v(w_1) + m_2 g_w(w_1)]g_b(w_2) \\ &+ [k_1 g_u(w_2) + l_1 g_v(w_2) + m_1 g_w(w_2)]g_b(w_1) \end{aligned} \tag{60}$$

Expanding the complex amplitudes and applying (57), the right hand side becomes

$$\begin{aligned} -i\nabla^2 A_0 z_1 z_2 e^{i\theta_1}e^{\theta_2} &= -i\nabla^2 A_0 z_1 z_2 e^{i\theta_0} \\ &= i \;\; k_0^2 + l_0^2 \;\; A_0 z_1 z_2 e^{i\theta_0}. \end{aligned} \tag{61}$$

The other terms on the right hand side of (26) can be expanded in a similar way:

$$-\partial_t\nabla^2(\mathbf{u}\cdot\nabla\mathbf{w}) = \;\; k_0^2 + l_0^2 \;\; \omega_0 B_0 z_1 z_2 e^{i\theta_0} \tag{62}$$

$$\partial_{zxt}(\mathbf{u}\cdot\nabla\mathbf{u}) = -m_0 k_0 \omega_0 C_0 z_1 z_2 e^{i\theta_0} \tag{63}$$

$$\partial_{zyt}(\mathbf{u}\cdot\nabla\mathbf{v}) = -m_0 l_0 \omega_0 D_0 z_1 z_2 e^{i\theta_0} \tag{64}$$

$$-f\partial_{zy}(\mathbf{u}\cdot\nabla\mathbf{u}) = if m_0 l_0 C_0 z_1 z_2 e^{i\theta_0} \tag{65}$$

$$f\partial_{zx}(\mathbf{u}\cdot\nabla\mathbf{v}) = -if m_0 k_0 D_0 z_1 z_2 e^{i\theta_0}. \tag{66}$$

Now setting the expanded left hand side (59) equal to the combined terms on the right hand side (61) - (66), and we can solve for the rate of change of the complex amplitude of wave 0:

$$\begin{aligned}\frac{\partial z_0}{\partial \tau} &= \frac{-i}{2\omega_0\kappa_0}\ \kappa_{H,0}^2(iA_0+\omega_0 B_0) - m_0\omega_0(k_0C_0+l_0D) + if m_0(l_0C_0-k_0D_0)\ \ z_1 z_2 \\ &=: -i\Gamma_0 z_1 z_2 \end{aligned} \tag{67}$$

where

$$\begin{aligned}\kappa_0^2 &= k_0^2+l_0^2+m_0^2 \\ \kappa_{H,0}^2 &= k_0^2+l_0^2.\end{aligned}$$

Equations for the modulation of waves 1 and 2 can be computed in a similar way, with a similar set of parameters $[A_1, B_1, C_1, D_1]$ and $[A_2, B_2, C_2, D_2]$ comprising the interaction coefficients Γ_1 and Γ_2.

References

[1] K. Polzin, J. Toole, J. Ledwell, and R. Schmitt, "Spatial variability of turbulent mixing in the abyssal ocean," Science **276**, 93 (1997).

[2] L. S. Laurent and C. Garrett, "The role of internal tides in mixing the deep ocean," J. Phys. Ocean. **32**, 2882 (2002).

[3] J. MacKinnon and K. Winters, "Tidal mixing hotspots governed by rapid parametric subharmonic instability," J. Phys. Ocean. **too appear**, (2004).

[4] D. Olbers and N. Pomphrey, "Disqualifying two candidates for the energy balance of oceanic internal waves," J. Phys. Ocean. **11**, 1423 (1981).

Elastic-skinned gravity currents

Anja Slim

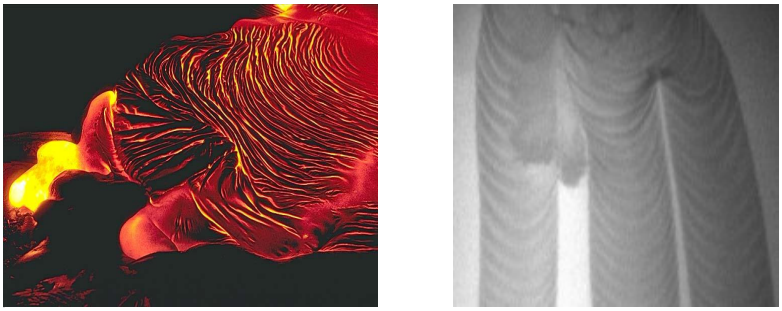
1 Introduction

We investigate a thin layer of very viscous fluid that flows down a slope due to gravity. However, instead of the usual free surface, the fluid is covered by an elastic plate. This allows for interesting wrinkling patterns to develop on the fluid surface.

One motivation for considering this problem is lava flows. A pahoehoe lava flow could be crudely considered an elastic-skinned gravity current. As a flow advances, it solidifies at the surface, forming a crust whilst remaining molten in the interior. An interesting aspect of these flows is the surface features that can be produced. One form, known as 'ropy' pahoehoe lava, is shown in figure 1(a). The traditional explanation for this phenomena is that the fluid beneath the crust pulls at it, and drags it into folds (see Ollier [2]). However, Fink and Fletcher [3] observed that the buckling was most prominent downstream of a constriction or a sudden decrease in bottom slope. They proposed that the explanation for buckling is that faster fluid pushes crust above it into crust above slower fluid. The resulting compression causes buckling. They also modelled this assuming the lava was semi-infinite and with a depth dependent viscosity.

A second motivation comes from experiments done by Clayton and Belmonte [1] with micellar fluid. Instead of mixing the two components of micellar fluid (aqueous solutions of surfactant and organic salt), they pumped one into the other. This resulted in a gravity current forming with an elastic interface between the two components. Clayton and Belmonte performed experiments on a slope, and observed fingering similar to that seen in a viscous fluid (see, for example, Huppert [4]). However when the volume flux was reduced during an experiment, buckling was observed along the fingers, as seen in figure 1(b).

The aim of this report is to begin to understand elastic-skinned gravity currents and buckling on their surfaces. It is organized as follows. In §2 we give a brief overview of previous work done on elastic-skinned gravity currents and general problems in hydro-elasticity. In §3 we outline some simple experiments and results. These experiments were performed to obtain a qualitative appreciation of the main features of these gravity currents. In §4 we formulate the governing system of equations. In §5 we present some preliminary analysis on three problems related to the currents. The first considers how the fluid influences buckling, the second considers how fluid traction at the base of the elastic plate can induce buckling and the final problem begins to combine these aspects to model the gravity current of the experiments in §3. Finally, §6 presents some conclusions and discusses future work.



(a) (b)

Figure 1: (a) Ropy pahoehoe lava. (b) Wrinkling observed on a gravity current finger between aqueous surfactant and organic salt solutions. *From* Clayton and Belmonte [1].

2 Previous work on hydro-elasticity

There has been much work done on problems of interaction of elastic materials with fluids. Major areas of application are to blood flow, flutter and drag reduction. Arteries are modelled as elastic-walled tubes filled with fluid, see for example Grotberg and Jensen [5]. Flutter usually refers to high Reynolds number flow over thin objects which causes vibrations, for example on aeroplanes (see, for example, Gee [6]). The application to drag reduction comes from an observation (for example Benjamin [7]) that replacing part of a solid wall by an elastic membrane increases the background velocity at which a turbulent boundary layer forms. Geophysical applications include modelling wave propagation through ice sheets on the ocean surface as waves in an elastic plate above an inviscid fluid (for example Chakrabarti et al. [8]). A further geophysical application is to magma flow through elastic-walled volcanic conduits (Balmforth et al. 2004 [9]).

To our knowledge, the only work that has investigated elastic-skinned gravity currents is the recent work of Hosoi and Mahadevan [10]. They studied the two-dimensional 'peeling' problem associated with fluid forced between an elastic plate and a rigid base, with which the plate is initially in contact. They model this system using the lubrication equation coupled with a linear plate equation.

3 Experiments

To investigate the behaviour of elastic-skinned gravity currents, we carried out experiments with a set-up as shown in figure 2. An elastic plate, either cling-wrap or exercise stretch band, was clamped to a rigid base and corn syrup was pumped between them.

The experiments were preliminary, although some qualitative observations can be made. Figure 3(a) shows the system with no initial slack (the plate was initially flat on the rigid base). Two important points about this configuration are that no wrinkles are visible and that the black line, which was perpendicular to the clamped edges before the fluid was pumped in, has not perceptively deformed. Figure 3(b) shows the system with an

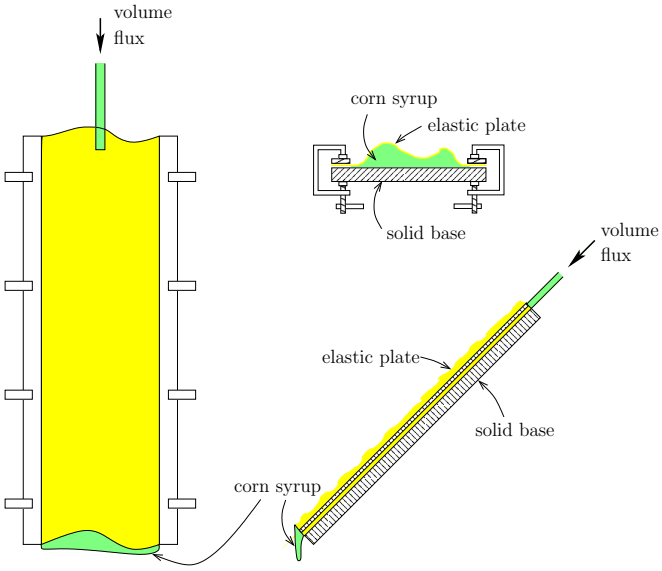
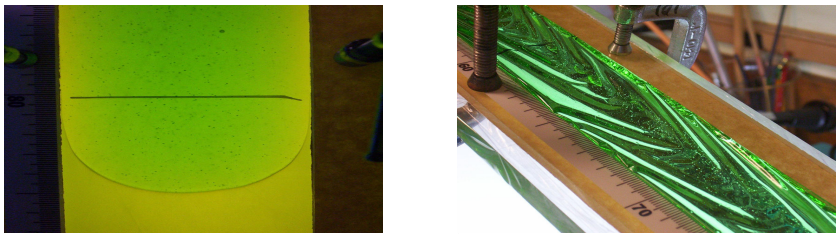


Figure 2: The experimental arrangement.

initial slack and with a fluid volume insufficient to fill it. Wrinkles are prominent and are clearly oriented downstream. Some other general points are: for a given flux and slack, the wavelength of the wrinkles increases with increasing width and increasing angle of slope; and the wavelength increases as the material becomes stiffer.[1]



(a) (b)

Figure 3: (a) Advancing front of the fluid beneath the elastic plate (exercise stretch band) for a case with no initial slack. No wrinkles are apparent. (b) A case with initial slack and clear wrinkles in the elastic plate (cling wrap).

[1]The wavelength was also observed to increase downstream from the source, but this is possibly an artefact of the way fluid was injected.

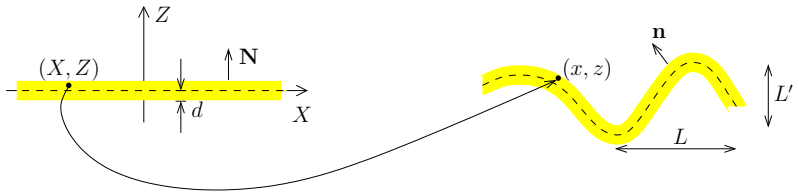


Figure 4: The elastic plate in its (left) undeformed and (right) deformed configuration. The undeformed coordinates are (X, Z) and the deformed (x, z). The horizontal length scale for the deformation of the plate is L and the vertical L'. The half-thickness of the plate is d.

4 Governing equations

There are three aspects of the configuration that need to be modelled: the elastic plate, the viscous fluid and the interaction between the two. We present equations for these in §§4.2, 4.3 and 4.4 respectively; together with boundary conditions on the plate and the fluid in §4.5. For the elastic plate we use the generalized von Kármán plate equations, which are nonlinear and include buckling terms. The fluid is described by the Stokes' equations.

Before presenting the governing equations for each component of the system, we give a derivation of the two-dimensional generalized von Kármán plate equations, to obtain an idea of the conditions under which they may be applied.

4.1 Derivation of the two-dimensional generalized von Kármán equations

The elastic plate is assumed to be very thin compared to the typical plate dimension, hence we derive the generalized two-dimensional von Kármán plate equations via asymptotics. By 'generalized', we refer to the particular form of the equations that have extra terms for in-plane forcings on the top and bottom of the plate: the original von Kármán equations only included normal forcing (see Yu [11]).

The usual method of derivation is from force balances. For example, Fung [12] derives the original von Kármán equations and Timoshenko [13] a generalized form with in-plane forcings on the plate (although some terms that are included in Yu and the derivation we present are not included there). An alternative method of derivation is a variational approach. For example, Landau and Lifschitz [14] derive the original equations and Yu the generalized.

Ciarlet [15] gave an asymptotic derivation for the original von Kármán equations. This was extended by various authors, for example, Millet et al. [16], to the generalized equations. In both cases, the authors assumed 'dead' loads: forcings that are independent of the deformation of the plate. Here we give a very similar derivation: our scalings are slightly different, but the expansion is essentially the same. We are also interested in forcings that do depend on the deformation of the plate to enable the coupling of the plate and the fluid.

We first give a brief overview of non-linear elasticity, to put our asymptotics into context. The system we are considering is shown in figure 4.

We may think of an elastic plate in two different coordinate systems. One is the undeformed, or Lagrangian, system in which everything is referred to the initial configuration

of the plate. The other is the deformed, or Eulerian, system in which things are referred to their current configuration. In two dimensions, the undeformed configuration is given by $\mathbf{X} = (X, Z)$ and the deformed configuration by $\mathbf{x} = (x, z)$ with

$$x = X + \xi(X, Z), \qquad z = Z + \zeta(X, Z), \tag{1}$$

where $\boldsymbol{\xi} = (\xi, \zeta)$ is the displacement.

The deformation gradient tensor converts between the two coordinate systems

$$F_{ij} = \partial x_i / \partial X_j.$$

The equation for conservation of momentum applies in the interior of the plate. From a fluid dynamics perspective, it may be considered most naturally in the deformed configuration and is given by

$$\nabla_{\mathbf{x}} \cdot \boldsymbol{\sigma} = 0, \tag{2}$$

where $\nabla_{\mathbf{x}}\cdot$ is the divergence and $\boldsymbol{\sigma}$ is the Cauchy stress tensor, both in the deformed coordinates. Throughout the report, subscripts $\mathbf{x}$ and $\mathbf{X}$ emphasize quantities in the deformed and undeformed configuration respectively. This equation assumes no body forces act on the plate and that the plate is in equilibrium. In our fluid–plate interaction problem elastic waves are fast compared to the viscous fluid flow and hence the plate at any instant may be assumed in equilibrium.

The boundary conditions on the plate are

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}_{\mathbf{x}} \tag{3}$$

on the upper and lower surfaces of the plate, $z = \pm d + \zeta(X, \pm d)$, where d is the constant half-thickness of the plate, $\mathbf{n}$ is the normal to the plate and $\mathbf{t}_{\mathbf{x}}$ is the surface traction imposed on it.

It is easier to derive the plate equations in the undeformed configuration because then the normal is simpler and we do not have to keep track of the location of the boundary of the plate. Although, in order to couple the plate equations with the fluid, we ultimately convert the derived equations back into the deformed coordinate system. Using the relation (see, for example, Antman [17] pg. 405, 406)

$$\boldsymbol{\sigma} = J^{-1}\mathbf{F}\boldsymbol{\Sigma}\mathbf{F}^T,$$

where $\boldsymbol{\Sigma}$ is the second Piola–Kirchoff stress tensor for the stress referred to the undeformed configuration, and $J = \det \boldsymbol{F}$, we translate (2) and (3) into the undeformed coordinate system as

$$\nabla_{\mathbf{X}} \cdot \boldsymbol{\Sigma}\mathbf{F}^T = 0, \tag{4}$$

$$\mathbf{F}\boldsymbol{\Sigma} \cdot \mathbf{N} = \sqrt{(1 + \xi_{,X})^2 + \zeta_{,X}^2}\,\mathbf{t}_{\mathbf{x}}, \tag{5}$$

where $\mathbf{N}$ is the normal in the undeformed configuration, $\mathbf{t}_{\mathbf{x}}$ is still the traction in the deformed configuration[2] and the boundary condition is now at $Z = \pm d$. The "$,\cdot$" subscript

[2]Note that this forcing still depends on $\boldsymbol{\xi}$. This means it should also be expanded and is something that still needs to be cleaned up.

refers to derivatives, a notation we use throughout the report. We introduce this notation to avoid confusion because later we will use plain subscripts, such as Σ_{XX}, to denote components of tensors.

The stress tensor, $\mathbf{\Sigma}$, may be related to the strain tensor, $\mathbf{E}$, through the linear, isotropic constitutive relation

$$\mathbf{\Sigma} = 2\tilde{\mu}\mathbf{E} + \tilde{\lambda}E_{kk}\mathbf{I}, \tag{6}$$

where $\tilde{\lambda}$, $\tilde{\mu}$ are Lamé constants and $\mathbf{E}$ is the non-linear strain tensor given by

$$E_{ij} = \frac{1}{2}\left(\frac{\partial \xi_i}{\partial X_j} + \frac{\partial \xi_j}{\partial X_i} + \frac{\partial \xi_k}{\partial X_i}\frac{\partial \xi_k}{\partial X_j}\right). \tag{7}$$

Note that using a linear constitutive model but non-linear strain is consistent provided that the strains are small (Stoker [18]).

We now look for solutions of the system (4) and (5) for which $\epsilon = L'/L \ll 1$, where L' is a length-scale associated with the out-of-plane displacements and L is a horizontal length-scale. Setting $L' = d$ is identical to Ciarlet [15] and Millet et al. [16]. We use the scaling

$$\xi = \epsilon L'\hat{\xi}, \quad \zeta = L'\hat{\zeta}, \quad X = L\hat{X}, \quad Z = \delta L'\hat{Z}, \quad t_x = \tilde{\mu}\epsilon^3\hat{t_x}, \quad t_z = \tilde{\mu}\epsilon^4\hat{t_z}, \tag{8}$$

where $\delta = d/L'$ and ϵ is much less than δ in a way made precise below. These assumptions are consistent with the scalings suggested in our experiments of §3. Scaling the momentum equation (4), neglecting hats, then gives

$$\Sigma_{XX} + \epsilon^2\xi_{,X}\Sigma_{XX} + \frac{1}{\delta}\epsilon\xi_{,Z}\Sigma_{XZ}\Big)_{,X} + \Big(\frac{1}{\epsilon}\frac{1}{\delta}\Sigma_{XZ} + \frac{1}{\delta^2}\xi_{,Z}\Sigma_{ZZ} + \epsilon\frac{1}{\delta}\xi_{,X}\Sigma_{XZ}\Big)_{,Z} = 0 \tag{9}$$

$$\zeta_{,X}\Sigma_{XX} + \epsilon\Sigma_{XZ} + \epsilon\frac{1}{\delta}\zeta_{,Z}\Sigma_{XZ}\Big)_{,X} + \Big(\frac{1}{\delta}\Sigma_{ZZ} + \frac{1}{\delta^2}\zeta_{,Z}\Sigma_{ZZ} + \epsilon\frac{1}{\delta}\zeta_{,X}\Sigma_{XZ}\Big)_{,Z} = 0, \tag{10}$$

and the boundary conditions (5) gives

$$\epsilon^{-1}\Sigma_{XZ} + \frac{1}{\delta}\xi_{,Z}\Sigma_{ZZ} + \epsilon\xi_{,X}\Sigma_{XZ} = \begin{cases} 0 & \text{on } Z = 1 \\ \tilde{\mu}\epsilon^2 t_x & \text{on } Z = -1 \end{cases}, \tag{11}$$

$$1 + \frac{1}{\delta}\zeta_{,Z}\Big)\Sigma_{ZZ} + \epsilon\zeta_{,X}\Sigma_{XZ} = \begin{cases} 0 & \text{on } Z = 1 \\ \tilde{\mu}\epsilon^4 t_z & \text{on } Z = -1 \end{cases}. \tag{12}$$

Into these equations, we substitute the formal expansion,

$$\xi = \xi_0(\mathbf{X}) + \epsilon^2\xi_2(\mathbf{X}) + \epsilon^4\xi_4(\mathbf{X}) + \cdots, \quad \zeta = \zeta_0(\mathbf{X}) + \epsilon^2\zeta_2(\mathbf{X}) + \epsilon^4\zeta_4(\mathbf{X}) + \cdots,$$

where only even order terms have been included, as odd order ones are irrelevant in the results we present.

Expanding $\mathbf{E}$ in powers of ϵ, from (7) and dropping the hats, gives

$$\begin{aligned} E_{XX} &= E^{(0)}_{XX} + \epsilon^2 E^{(2)}_{XX} + \cdots \\ &= 0 + \epsilon^2 \quad \xi_{0,X} + \frac{1}{2}\zeta^2_{0,X} \quad + \cdots \\ E_{ZZ} &= E^{(0)}_{ZZ} + \epsilon^2 E^{(2)}_{ZZ} + \cdots \\ &= \quad \delta^{-1}\zeta_{0,Z} + \frac{1}{2}\delta^{-2}\zeta^2_{0,Z} \quad + \epsilon^2 \quad \delta^{-1}\zeta_{2,Z} + \frac{1}{2}\delta^{-2}\xi^2_{0,Z} + \delta^{-2}\zeta_{0,Z}\zeta_{2,Z} \quad + \cdots \end{aligned}$$

for the normal strain components and

$$\begin{aligned} E_{XZ} &= \epsilon E^{(1)}_{XZ} + \cdots \\ &= \epsilon \quad \frac{1}{2}\delta^{-1}\xi_{0,Z} + \frac{1}{2}\zeta_{0,X} + \frac{1}{2}\delta^{-1}\zeta_{0,X}\zeta_{0,Z} \quad + \cdots \end{aligned}$$

for the shear strain. From these, the stresses may be written, using (6) as

$$\Sigma^{(k)}_{XX} = (\tilde{\lambda} + 2\tilde{\mu})E^{(k)}_{XX} + \tilde{\lambda}E^{(k)}_{ZZ}, \quad \Sigma^{(k)}_{XZ} = 2\tilde{\mu}E^{(k)}_{XZ}, \quad \Sigma^{(k)}_{ZZ} = \tilde{\lambda}E^{(k)}_{XX} + (\tilde{\lambda} + 2\tilde{\mu})E^{(k)}_{ZZ}.$$

We can now begin with the asymptotics.

At order 1, the equations and boundary conditions (9), (11) and (10), (12) imply

$$\Sigma^{(0)}_{ZZ} \quad 1 + \delta^{-1}\zeta_{0,Z} \quad = 0, \quad \Sigma^{(1)}_{XZ} + \delta^{-1}\xi_{0,Z}\Sigma^{(0)}_{XX} = 0,$$

respectively. The former of these gives three possible solutions and we choose the physically most meaningful

$$\Sigma^{(0)}_{ZZ} = \zeta_{0,Z} = 0.$$

Hence to leading order the out-of-plane displacement is independent of Z. The second equation then gives

$$\Sigma^{(1)}_{XZ} = 0, \qquad \xi_0(X, Z) = -\delta\zeta_{0,X}Z + \tilde{\xi}_0(X),$$

where $\tilde{\xi}_0(X)$ may be interpreted as the in-plane displacement of the centre-line of the plate.

At order ϵ^2, the equations become

$$\Sigma^{(2)}_{XX,X} + \delta^{-1}\Sigma^{(3)}_{XZ,Z} = 0, \tag{13}$$

subject to $\Sigma^{(3)}_{XZ} = 0$ on $Z = 1$ and $\Sigma^{(3)}_{XZ} = \tilde{\mu}t_z$ on $Z = -1$ from (9) and (11), and from (10) and (12)

$$\Sigma^{(2)}_{ZZ} = 0.$$

This last equation may be solved for ζ_2 to give

$$\zeta_2 = \frac{\tilde{\lambda}}{\tilde{\lambda} + 2\tilde{\mu}} \quad \delta^2\zeta_{0,XX}\frac{1}{2}Z^2 - \delta\tilde{\xi}_{0,X}Z \quad - \delta\frac{\tilde{\lambda} + \tilde{\mu}}{\tilde{\lambda} + 2\tilde{\mu}}\zeta^2_{0,X},$$

where constant terms of integration have been neglected because they are not relevant in our derivation. Equation (13) gives

$$\Sigma^{(3)}_{XZ} = -\tilde{\mu}\delta\frac{4(\tilde{\lambda}+\tilde{\mu})}{\tilde{\lambda}+2\tilde{\mu}} \quad -\delta\zeta_{0,XXX}\frac{1}{2}Z^2 + \tilde{\xi}_{0,XX}Z + \frac{1}{2}\ \zeta^2_{0,X}{}_{\ X}Z \ \ + A(X),$$

where $A(X)$ is a constant function (in Z) of integration. $A(X)$ may be eliminated using $\left[\Sigma^{(3)}_{XZ}\right]^1_{-1} = -\tilde{\mu}t_x$ from the boundary conditions, to give the first of the generalized von Kármán equations:

$$\frac{8(\tilde{\lambda}+\tilde{\mu})}{\tilde{\lambda}+2\tilde{\mu}}\delta\ \ \tilde{\xi}_{0,XX} + \frac{1}{2}\ \ \zeta^2_{0,X}{}_{\ X} \ \ = t_x. \tag{14}$$

This equation governs the in-plane deformation of the plate. We will also require the solution for $A(X)$, which may be found by taking the sum $\Sigma^{(3)}_{XZ}\Big|_{Z=1} + \Sigma^{(3)}_{XZ}\Big|_{Z=-1}$ and is

$$2A(X) = \tilde{\mu}t_x - \tilde{\mu}\delta^2\frac{4(\tilde{\lambda}+\tilde{\mu})}{\tilde{\lambda}+2\tilde{\mu}}\zeta_{0,XXX}.$$

At order ϵ^4, equations (10) and (12) combine to give

$$\int_{-1}^{1}\ \zeta_{0,X}\Sigma^{(2)}_{XX} + \Sigma^{(3)}_{XZ}\ _X\,\mathrm{d}Z = -\delta^{-1}\left[\zeta_{0,X}\Sigma^{(3)}_{XZ} + \Sigma^{(4)}_{ZZ}\right]^1_{-1} = -\delta^{-1}\tilde{\mu}t_z,$$

from which, provided $\delta^3 \gg \epsilon$, the second generalized von Kármán equation becomes

$$\frac{8(\tilde{\lambda}+\tilde{\mu})}{3(\tilde{\lambda}+2\tilde{\mu})}\delta^3\zeta_{0,XXXX} = -t_z + \frac{8(\tilde{\lambda}+\tilde{\mu})}{\tilde{\lambda}+2\tilde{\mu}}\delta\ \ \zeta_{0,X}\tilde{\xi}_{0,X} + \frac{1}{2}\zeta^3_{0,X}{}_{\ X} + \delta t_{x,X}, \tag{15}$$

which governs the out-of-plane flexing of the plate. If $\delta^3 \sim \epsilon$, then the first term, the flexural term, should be neglected and we obtain a membrane model. The previous orders appear to be consistent in the asymptotic expansion provided $\delta^2 \gg \epsilon$.

We now transform equations (14) and (15) back into the deformed coordinate system. From (1) and the scalings (8), we have

$$\frac{\partial}{\partial X} = \frac{\partial}{\partial x} + \mathcal{O}(\epsilon).$$

Hence, by replacing all derivatives in X by derivatives in x, we obtain the generalized von Kármán equations, on rescaling, in the deformed configuration.

We can also calculate the energy density for the elastic plate from the asymptotics. For an elastic body, the energy density becomes

$$\psi = \frac{1}{2}\Sigma_{ij}E_{ij} = \epsilon^4\ \ \frac{1}{2}(\tilde{\lambda}+2\tilde{\mu})\ \ E^{(2)2}_{XX} + E^{(2)2}_{ZZ}\ \ + \tilde{\lambda}E^{(2)}_{XX}E^{(2)}_{ZZ}\ \ + \mathcal{O}(\epsilon^6) \tag{16}$$

$$= \epsilon^4\ \frac{4(\tilde{\lambda}+\tilde{\mu})\tilde{\mu}}{\tilde{\lambda}+2\tilde{\mu}}\ \ \frac{1}{3}\delta^2\zeta^2_{0,XX} + \ \ \tilde{\xi}_{0,X} + \frac{1}{2}\zeta^2_{0,X}\ \ \ + \mathcal{O}(\epsilon^6). \tag{17}$$

These equations, on rescaling, become the two-dimensional version of the energy equation (see, for example, Landau and Lifschitz pg. 58 [14]).

4.2 The three-dimensional generalized von Kármán equations

In three dimensions, the equations may be decomposed as follows.

For a displacement (ξ, η, ζ), the in-plane strains are given by

$$e_{xx} = \xi_x + \frac{1}{2}\zeta_{,x}^2, \quad e_{yy} = \eta_{,y} + \frac{1}{2}\zeta_{,y}^2, \quad e_{xy} = \frac{1}{2}\left(\xi_{,y} + \eta_{,x} + \zeta_{,x}\zeta_{,y}\right), \tag{18}$$

from which the in-plane stresses may be obtained

$$N_x = \frac{2d\mathcal{E}}{1-\nu^2}(e_{xx} + \nu e_{yy}), \quad N_y = \frac{2d\mathcal{E}}{1-\nu^2}(\nu e_{xx} + e_{yy}), \quad N_{xy} = \frac{2d\mathcal{E}}{1+\nu}e_{xy}, \tag{19}$$

where $\mathcal{E} = \tilde{\mu}(3\tilde{\lambda} + 2\tilde{\mu})/(\tilde{\lambda} + \tilde{\mu})$ is Young's modulus, $\nu = \tilde{\lambda}/2(\tilde{\lambda} + \tilde{\mu})$ is the Poisson ratio and d is the half-thickness of the plate.

This gives the in-plane part of the generalized von Kármán equations as

$$N_{x,x} + N_{xy,y} = t_x, \quad N_{xy,x} + N_{y,y} = t_y \tag{20}$$

and the flexural part of the generalized von Kármán equations is described by

$$B\nabla_H^4\zeta = -t_z + (N_x\zeta_{,x} + N_{xy}\zeta_{,y})_{,x} + (N_{xy}\zeta_{,x} + N_y\zeta_{,y})_{,y} + dt_{x,x} + dt_{y,y}, \tag{21}$$

where $B = 2d^3\mathcal{E}/3(1-\nu^2)$ is the flexural rigidity of the material, (t_x, t_y, t_z) is the traction on the base of the plate and $\nabla_H^4 = (\partial_x^2 + \partial_y^2)^2$.

4.3 Fluid equations

The fluid is assumed very viscous and so the Stokes' equations are used as the governing equations with

$$\mu\nabla^2\mathbf{u} = \boldsymbol{\nabla} p + \rho g(-\sin\theta, 0, \cos\theta), \quad \nabla\cdot\mathbf{u} = 0, \tag{22}$$

where $\mathbf{u} = (u, v, w)$ is the fluid velocity, p pressure, μ fluid viscosity, ρ fluid density, g gravity and θ the angle of the slope of the system. The x-axis is assumed to point downslope, y across slope and z perpendicular to the slope.

4.4 Interface conditions

At the interface between the plate and the fluid, we require continuity of velocity

$$\left.\frac{\partial\boldsymbol{\xi}}{\partial t}\right|_{\mathbf{X}} = \frac{\partial\boldsymbol{\xi}}{\partial t} + \mathbf{u}\cdot\nabla_{\mathbf{x}}\boldsymbol{\xi} = \mathbf{u}. \tag{23}$$

It should be noted that this is the only place where there is time dependence in the governing system of equations.

The tractions on the top of the plate are taken to be zero, and those on the bottom of the plate are

$$t_x = \hat{\mathbf{t}}_1\cdot\boldsymbol{\sigma}_f\cdot\hat{\mathbf{n}}, \quad t_y = \hat{\mathbf{t}}_2\cdot\boldsymbol{\sigma}_f\cdot\hat{\mathbf{n}}, \quad t_z = \hat{\mathbf{n}}\cdot\boldsymbol{\sigma}_f\cdot\hat{\mathbf{n}},$$

where $\boldsymbol{\sigma}_f$ is the fluid stress at the plate boundary, $\hat{\mathbf{n}}$ is the unit normal to the (base of the) plate and $\hat{\mathbf{t}}_1$, $\hat{\mathbf{t}}_2$ are the tangents to the plate aligned with x- and y-axis respectively. To as many terms as we need in our problems, these tractions become

$$t_x = \mu(u_{,z} + w_{,x}) - 2\mu(u_{,x} - w_{,z})\zeta_{,x} - \mu(u_{,y} + v_{,x})\zeta_{,y} \tag{24}$$

$$t_y = \mu(v_{,z} + w_{,y}) - 2\mu(v_{,y} - w_{,z})\zeta_{,y} - \mu(u_{,y} + v_{,x})\zeta_{,x} \tag{25}$$

$$t_z = -p + 2\mu w_{,z} - 2\mu(u_{,z} + w_{,x})\zeta_{,x} - 2\mu(v_{,z} + w_{,y})\zeta_{,y}. \tag{26}$$

4.5 Boundary conditions

The boundary conditions on the plate edges are given by fixed displacement, usually

$$\xi = \eta = \zeta = 0 \tag{27}$$

and a further condition on ζ, since the flexural equation (21) is fourth order. In our problems, the edges of the plate are at $y = \text{const.}$ and possibilities for the further boundary condition on ζ include $\zeta_{,y} = 0$, $\zeta_{,yy} = 0$, $\zeta_{,yyy} = 0$ for clamped, freely-supported or hinged respectively. For all the problems we study, we use a clamped boundary condition at the edges:

$$\zeta_{,y} = 0. \tag{28}$$

The fluid has zero velocity on rigid surfaces.

5 The problems

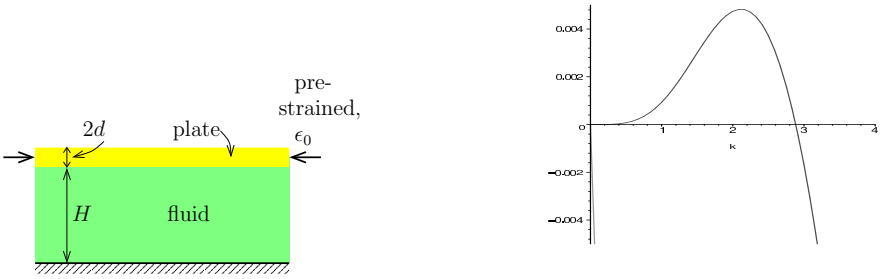
In this section, we consider three problems of fluid–plate interaction. The first problem, in §5.1, is the buckling of a compressed beam in contact with a two-dimensional very viscous fluid below it. This is a modification of the classical, purely elastic problem of beam buckling. It was first studied by Huang and Suo [19] using the lubrication equation to model the fluid. The motivation for considering this problem is to understand how the fluid modifies buckling.

The second problem, in §5.2, looks at flow in a channel down a slope. In this case there is no pre-compression on the system and the aim is to understand under what conditions the fluid induced shear on the plate can cause buckling.

Finally, in §5.3, we bring aspects from both of the previous problems together to try to begin to understand the wrinkles observed in our experiments of §3.

We use linear stability analysis to analyze our problems. This allows us to find when and where wrinkling can occur and is a relatively simple method. Linear stability analysis is not particularly relevant to pure elasticity cases because elastic waves are fast, and so the system quickly passes the linear growth regime. However, for a viscous fluid controlling the growth rate of any buckling instabilities in our problems, a linear stability analysis is more useful.

Our notation throughout is subscript 0 for base states and no subscripts for the perturbation flow. Derivatives are denoted as in §4 by subscript "$,\cdot$" for consistency.



(a) (b)

Figure 5: (a) The configuration for the buckling of a beam above a very viscous fluid. (b) Growth rate, $\Im(\omega)$, versus wavenumber k. Note that one mode is very close to the negative $\Im(\omega)$-axis.

5.1 Two-dimensional simple buckling

In this subsection, we examine the buckling of a pre-compressed beam in contact with a two-dimensional very viscous fluid of depth H, as shown in figure 5(a). This problem was first investigated by Huang and Suo [19] in the context of the manufacture of semi-conductors. Our problem formalism is very similar to theirs, however we use the Stokes' equations for the fluid whereas they used the lubrication equation. Our results are essentially identical to theirs. In addition, Huang and Suo also performed a numerical simulation of the system. For this problem, gravity is ignored and the fluid is passive.

5.1.1 Base state

In the base state, everything is assumed to be zero, except for the normal strain e_{xx} which is non-zero. We refer to this as pre-strain, $e_{xx} = \epsilon_0$. Hence the system has a given pre-stress: compressive if $\epsilon_0 < 0$ and tensile if $\epsilon_0 > 0$.

5.1.2 Linearized perturbation

We now add a perturbation and linearize the governing system of equations (18)–(28) about the base state. In addition, we non-dimensionalize the equations using $L = H$ for lengths, $U = 2d\mathcal{E}/(1-\nu^2)\mu$ for velocities, L/U for times and $\mu U/L$ for pressures. Note that the time scale of the problem depends on the fluid. We set $\delta_b = d/H$.

The governing equations then reduce to

$$\begin{aligned} \xi_{,xx} = u_{,z} + w_{,x}, \quad & \frac{1}{3}\delta_b^2 \zeta_{,xxxx} = p - 2w_{,z} + \epsilon_0 \zeta_{,xx}, \\ \nabla^2 \mathbf{u} = \boldsymbol{\nabla} p, \quad & \nabla \cdot \mathbf{u} = 0, \\ \xi_{,t} = u, \quad & \zeta_{,t} = w \quad \text{on } z = 1, \\ \mathbf{u} = 0 \quad & \text{on } z = 0, \end{aligned}$$

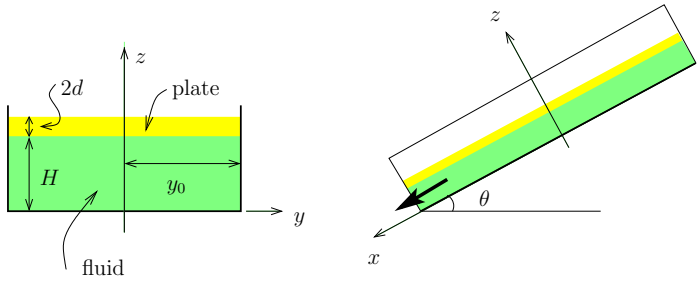


Figure 6: Configuration for channel flow. The fluid is confined by rigid walls at $y = \pm y_0$, and $z = 0$ and an elastic plate at $z = H$. The fluid flows due to gravity alone.

for the plate equations, fluid equations, interface conditions and boundary conditions respectively.

If we now look for solutions proportional to $e^{ikx-i\omega t}$, these equations reduce to

$$\det \begin{pmatrix} \sinh k + k\cosh k + 2\Omega(\cosh k + k\sinh k) & -k^2\sinh k - 2\Omega k^2\cosh k \\ \frac{1}{3}\delta_b^2 k^2 + \epsilon_0 \;\; \sinh k & \frac{1}{3}\delta_b^2 k^2 + \epsilon_0 \;\; (\sinh k - k\cosh k) \\ -\delta_b(\sinh k + k\cosh k) + 2\Omega\cosh k & +\delta_b k^2\sinh k + 2\Omega(\cosh k - k\sinh k) \end{pmatrix} = 0,$$

where $\Omega = -i\omega/k$, to give the dispersion relation. Figure 5(b) shows the two solution branches for $\Im(\omega)$ against k, for the specific choice of parameters $\epsilon_0 = -0.0156$, $\delta_b = 0.075$. As can be seen, one root is negative for all k and hence the perturbation is stable, whereas the other is positive for small k, and hence the perturbation is unstable. The short-wave cut-off occurs at $\delta_b^2 k^2/3 + \epsilon_0 = 0$ and is due to flexural rigidity preventing shorter waves.

The main contribution of the fluid is setting the time-scale: its presence reduces the rate at which the instability grows, because it only slowly flows out of the way of the buckling plate (Huang and Suo).

5.2 Channel flow

In this subsection we examine the flow of a very viscous fluid in a channel of half-width y_0, down a slope of angle θ. The side walls and base of the channel are rigid and the top, at $z = H$, is an elastic plate, clamped at its edges across the slope. The fluid flow is due only to gravity: there is no imposed pressure gradient. It is assumed that the volume flux of fluid is exactly sufficient to keep the elastic plate flat in its base state. This configuration is shown in figure 6.

This is closely related to a much studied classical problem in elasticity: a thin plate in shear (see, for example, Wong and Pellegrino [20]). In this case the configuration is a plate whose long edges are attached to rigid rods held a fixed distance apart, so that the plate is just taut. One of the rods is then translated parallel to its length. Wrinkles are observed between the rods, at 45°, regardless of the amplitude of the displacement of the rod. This is a result of the Poisson effect: although the plate is under tension in the direction of the major principal stress, it is in compression in the direction of minor principal stress with the

result that buckling occurs. This problem has been approached in a variety of ways. The original method was using tension field theory, see for example Mansfield [21], which was able to predict the 45° angle of the wrinkle.[3] More recently, Wong and Pellegrino [22] found simple analytic predictions using an assumed form for the out-of-plane displacement and force or energy minimization arguments. In addition, experiments and many finite element computations have been performed (see, for example Wong and Pellegrino [20]).

The aim of this section is to show that, in our problem, the traction of the fluid on the bottom of the elastic plate, results in a similar buckling phenomenon.

5.2.1 Base flow

For the base state, we assume that the fluid flow is purely down-slope with $v_0 = w_0 = 0$. We assume that the plate is not displaced in the cross-slope direction, so $\eta_0 = 0$, and that it suffers no out-of-plane displacement, hence $\zeta_0 = H$. The system is assumed independent of the downslope coordinate, x, and time, t. Hence we take

$$u = u_0(y, z), \qquad \xi = \xi_0(y), \qquad p = p_0(z).$$

The governing equations for the fluid (22) in the base state become

$$\mu(u_{0,yy} + u_{0,zz}) = -\rho g \sin\theta, \quad 0 = p_{0,z} + \rho g \cos\theta, \tag{29}$$

with boundary conditions

$$u_0 = 0 \qquad \text{on } z = 0, H;\ y = \pm y_0. \tag{30}$$

This system is identical to that for flow in a rigid rectangular duct: the fact that the upper surface is elastic is immaterial (except for the pressure being known at the upper surface, see below).

The in-plane deformation equations for the plate (20) reduce to

$$\frac{d\mathcal{E}}{1+\nu}\xi_{0,yy} = \mu u_{0,z}, \tag{31}$$

with boundary condition $\xi_0 = 0$ at $y = \pm y_0$. The flexure equation (21) reduces to $p_0 = 0$ on $z = H$, completing the pressure equation (29b).

Using Fourier series for u_0, this system can be solved giving

$$p_0 = \rho g \cos\theta\,(H - z), \tag{32}$$

$$u_0 = \rho g \sin\theta\, H^2/\mu\, \mathcal{U}(Y, Z; \alpha), \tag{33}$$

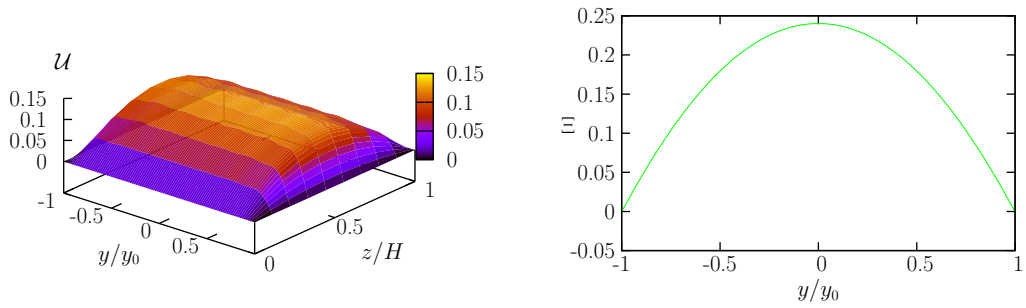
$$\xi_0 = y_0^2 \rho g \sin\theta\, H(1+\nu)/d\mathcal{E}\ \Xi(Y; \alpha), \tag{34}$$

where $Y = y/y_0$, $Z = z/H$, $\alpha = H/y_0$ and

$$\mathcal{U}(Y, Z; \alpha) = \frac{16}{\pi^4} \sum_{n,m=0}^{\infty} \frac{(-1)^m \sin[(2n+1)\pi Z] \cos[(2m+1)\pi Y/2]}{(2n+1)(2m+1)[(2n+1)^2 + \alpha^2(2m+1)^2/4]},$$

$$\Xi(Y; \alpha) = \frac{64}{\pi^5} \sum_{n,m=0}^{\infty} \frac{(-1)^m \cos[(2m+1)\pi Y/2]}{(2m+1)^3[(2n+1)^2 + \alpha^2(2m+1)^2/4]}.$$

[3] Unfortunately, we cannot use tension field theory, as it relies on the existence of an Airy stress function: one no longer exists for the generalized von Kármán equations.



(a) (b)

Figure 7: (a) Velocity profile, $\mathcal{U}$, against $Y = y/y_0$ and $Z = z/H$ and (b) downstream displacement of the plate, Ξ, against Y for $\alpha = 0.25$.

Plots of solutions for $\mathcal{U}$ and Ξ are shown in figure 7 for $\alpha = 0.25$. As can be clearly seen, the fluid pulls the plate downstream.

To be compatible with the scalings assumed to obtain the generalized von Kármán plate equations, we must be in a regime in which the displacement is small. For this we use the strict scaling $\xi \sim d^2/L$ and hence we require

$$\mathfrak{E}_c = \frac{\mathcal{E} d^3}{\rho g \sin\theta \, H y_0^3} \sim 1. \tag{35}$$

5.2.2 Linearized perturbation

We now add a small perturbation to the base state and consider the linearized equations governing the system.

The linearized perturbation equations for the plate then become

$$\frac{2d\mathcal{E}}{1-\nu^2}(\xi_{,xx} + \nu\eta_{,xy}) + \frac{d\mathcal{E}}{1+\nu}(\xi_{,yy} + \eta_{,xy}) = \mu u_{0,zz}\zeta - \mu u_{0,y}\zeta_{,y} + \mu(u_{,z} + w_{,x}), \tag{36}$$

$$\frac{d\mathcal{E}}{1+\nu}(\xi_{,xy} + \eta_{,yy}) + \frac{2d\mathcal{E}}{1-\nu^2}(\nu\xi_{,xy} + \eta_{,yy}) = -\mu u_{0,y}\zeta_{,x} + \mu(v_{,z} + w_{,y}), \tag{37}$$

for the in-plane deformation equations (20) and

$$B\nabla_H^4\zeta = p_{0,z}\zeta + 2\mu u_{0,z}\zeta_{,x} + p - 2\mu w_{,z} + 2\frac{d\mathcal{E}}{1+\nu}\xi_{0,y}\zeta_{,xy} + \frac{d\mathcal{E}}{1+\nu}\xi_{0,yy}\zeta_{,x}$$
$$+ d[-\mu u_{0,y}\zeta_{,xy} + \mu u_{0,zz}\zeta_{,x} + \mu(u_{,xz} + w_{,xx}) - \mu u_{0,yy}\zeta_{,x} - \mu u_{0,y}\zeta_{,xy} + \mu(v_{,yz} + w_{,yy})], \tag{38}$$

for the flexural equation (21). The boundary conditions for the plate, from (27) and (28), are

$$\xi = \eta = \zeta = \zeta_{,y} = 0 \qquad \text{at } y = \pm y_0. \tag{39}$$

The linearized perturbation equations for the fluid, from (22), become

$$\mu\nabla^2 u = p_{,x}, \quad \mu\nabla^2 v = p_{,y}, \quad \mu\nabla^2 w = p_{,z}, \quad u_{,x} + v_{,y} + w_{,z} = 0, \tag{40}$$

with boundary conditions

$$u = v = w = 0 \qquad \text{on } z = 0;\ y = \pm y_0. \tag{41}$$

The condition of continuity of velocity at the interface (23) becomes

$$\xi_{,t} + v\xi_{0,y} = u + u_{0,z}\zeta, \quad \eta_{,t} = v, \quad \zeta_{,t} = w. \tag{42}$$

These equations appear fairly complex, however we can make some significant simplifications by neglecting terms which are order d/L smaller than the leading order terms (in fact, we must do this as it not consistent to keep them). Here L is the horizontal length scale for the perturbation, assumed to satisfy $L \ll y_0$. We scale according to (8) in the derivation and bear in mind the different length-scale for the base state. We also scale vertical derivatives with L as we would like vertical and horizontal length scales to be similar. We then have from $t_z \sim \mathcal{E}d^4/L^4$ in (8), as $\mathcal{E} \sim \mu$, that $p \sim \mathcal{E}d^4/L^4$, from which $u \sim \mathcal{E}/\mu\ d^4/L^3$ using (40). Substituting these into (36)–(38) and (42) we obtain the reduced system

$$\frac{2d\mathcal{E}}{1-\nu^2}(\xi_{,xx} + \nu\eta_{,xy}) + \frac{d\mathcal{E}}{1+\nu}(\xi_{,yy} + \eta_{,xy}) = \mu u_{0,zz}\zeta, \tag{43}$$

$$\frac{d\mathcal{E}}{1+\nu}(\xi_{,xy} + \eta_{,yy}) + \frac{2d\mathcal{E}}{1-\nu^2}(\nu\xi_{,xy} + \eta_{,yy}) = 0 \tag{44}$$

for the in-plane equations,

$$B\nabla_H^4\zeta = p_{0,z}\zeta + 2\mu u_{0,z}\zeta_{,x} + p - 2\mu w_{,z} + 2\frac{d\mathcal{E}}{1+\nu}\xi_{0,y}\zeta_{,xy} + \frac{d\mathcal{E}}{1+\nu}\xi_{0,yy}\zeta_{,x} \tag{45}$$

for the flexural equation and

$$0 = u + u_{0,z}\zeta, \quad 0 = v, \quad \zeta_{,t} = w \quad \text{on } z = H \tag{46}$$

for the interface conditions. The fluid remains as in (40) and boundary conditions remain (39) and (41).

To gain an understanding of the instability, we look for small wavelength solutions and we assume the background variation is slow in comparison. This allows us to Fourier decompose the perturbation in both the x- and the y-direction. Ultimately, this could be made rigorous using WKB.

Therefore we assume the perturbation is proportional to $e^{ikx+ily-i\omega t}$ and we ignore cross-stream boundary conditions.

Solving (40) subject to $u = v = w = 0$ on $z = 0$, and (46) on $z = H$, we obtain

$$p - 2\mu w_{,z} = 2\mu K i\omega \frac{\cosh KH \sinh KH + KH}{\sinh^2 KH - K^2H^2}\zeta + u_{0,z}\frac{2\mu ikK^2H^2}{\sinh^2 KH - K^2H^2}\zeta,$$

where $K^2 = k^2 + l^2$. Hence (45) becomes

$$\begin{aligned} BK^4\zeta = {} & -\rho g \sin\theta\,\zeta + 2\mu K i\omega \frac{\cosh KH \sinh KH + KH}{\sinh^2 KH - K^2H^2}\zeta + 2\mu iku_{0,z}\frac{\sinh^2 KH}{\sinh^2 KH - K^2H^2}\zeta \\ & - \frac{2d\mathcal{E}}{1+\nu}\xi_{0,y}kl\zeta + \frac{d\mathcal{E}}{1+\nu}ik\xi_{0,yy}\zeta. \end{aligned}$$

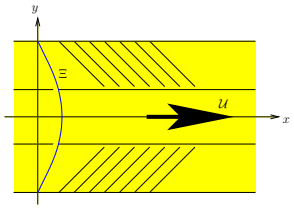


Figure 8: Cartoon (plan view) of the observed instability for the channel flow. The angled lines represent wrinkle crests. The interior area is wrinkle free.

This relation only depends on ζ and hence the dispersion relation becomes

$$-2\mu K\frac{\cosh KH\sinh KH+KH}{\sinh^2 KH-K^2H^2}i\omega=-BK^4-\rho g\cos\theta$$
$$-2y_0\rho g\sin\theta\, H\,\Xi'(Y;\alpha)kl+ik\frac{3\sinh^2 KH-K^2H^2}{\sinh^2 KH-K^2H^2}\rho g\sin\theta\, H\Xi''(Y;\alpha), \qquad (47)$$

using (31) and (34).

From the dispersion relation some of the properties of the instability may be deduced. We consider $\Im(\omega)$, and note that the coefficient of $i\omega$ is positive for all (k,l). The first two terms on the right hand side of (47) result from the flexural term and hydrostatic pressure. Both are always negative and hence stabilizing. The third term results from the base displacement of the plate due to fluid traction on the base and can be destabilizing. A necessary condition for instability is $kl\Xi'(Y;\alpha)<0$. Hence any observed crests will be oriented downstream, as shown in figure 8. In the interior $\Xi'=0$ and close by it is too small to overcome the stabilizing flexural and hydrostatic pressure terms. Hence, in an interior region, no wrinkles are observed. The final term of (47) is imaginary and hence not directly important to the stability.

Finally, if we look at a fastest growing mode, then the observed modes have $k=\pm l$. This suggests that even for a small displacement due to fluid traction, the wrinkles will have a significant angle.

There is much that needs to be done to complete this analysis, and check for its consistency with scalings.

5.3 Rivulet flow

In this subsection we study the flow of a very viscous fluid, contained between a rigid base and an elastic plate clamped to the base along lines $y=\pm y_0$. This configuration is shown in figure 9. It is close to the physical situation we wish to model and is equivalent to our experiments of §3. In this scenario, there is a combination of pre-strain, as in §5.1, and fluid traction on the base of the plate, as in §5.2, to produce buckling; bringing together the features of the previous two problems.

First consider the elastic plate in the absence of the fluid. It has approximate unstressed length $2(1-\epsilon_0)y_0$, or equivalently a pre-strain $e_{yy}=\epsilon_0$, where $|\epsilon_0|\ll 1$. If $\epsilon_0>0$, then the plate is in tension and is flat against the base. If $\epsilon_0<0$, then the plate is in compression and will buckle. Hence this pre-compression may be conceptualized as slack.

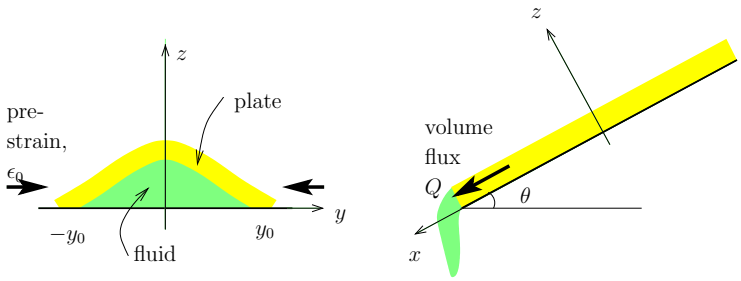


Figure 9: Configuration for the rivulet problem. An elastic plate is clamped to a rigid base along lines $y = \pm y_0$ and fluid is introduced into the gap between.

If fluid is now introduced between the rigid base and the taut plate for $\epsilon_0 > 0$, then it forces the plate out of the plane of the base, and increases the tension on it, and no instability is found. For $\epsilon_0 < 0$, the situation is more complex. If sufficient fluid is introduced, then the increased out-of-plane displacement is sufficient for the plate to be taut. However, if a small volume of fluid is introduced, some compressive sress may remain and the plate still buckles. In any case, the plate is forced out of the plane of the base by the fluid, stretched across stream to accommodate the fluid and pulled downstream by the fluid.

5.3.1 Base flow

We use the governing system of equations (18)–(28), non-dimensionalizing the variables by y_0 for lengths, $\rho g \cos\theta\, y_0$ for pressures and elastic stresses and $\rho g \sin\theta\, y_0^2/\mu$ for velocities. We also set $\mathfrak{E}_r = 2d\mathcal{E}/(1-\nu^2)\rho g \cos\theta\, y_0^2$, representing a ratio of elastic stress to fluid pressure, and put $\delta_r = d/y_0$.

The base state is assumed independent of down-stream coordinate, x, and $v_0 = w_0 = 0$ is assumed.

The in-plane plate equations (20) reduce to

$$\frac{1}{2}(1-\nu)\mathfrak{E}_r \xi_{0,yy} = u_{0,z} - u_{0,y}\zeta_{0,y}, \quad \mathfrak{E}_r \quad \epsilon_0 + \eta_{0,y} + \frac{1}{2}\zeta_{0,y}^2 \quad = \text{const.} = \mathfrak{E}_r N_0, \tag{48}$$

where ϵ_0 is the imposed pre-strain and the dependent variable N_0 may be interpreted as the remaining pre-stress. Note that, unlike in Euler's elastica problem, the stress is a dependent variable whereas the displacement is a parameter. The flexure equation (21) becomes

$$\frac{1}{3}\delta_r^2 \mathfrak{E}_r \zeta_{0,yyyy} = p_0 + \mathfrak{E}_r N_0 \zeta_{0,yy}, \tag{49}$$

and the boundary conditions on the plate are

$$\xi_0 = \eta_0 = \zeta_0 = \zeta_{0,y} = 0 \qquad \text{at } y = \pm 1. \tag{50}$$

The fluid equations (22) become

$$u_{0,yy} + u_{0,zz} = -1 \tag{51}$$

for the downstream flow and

$$p_0 = \tilde{p}_0 - z \tag{52}$$

for hydrostatic pressure, where $\tilde{p}_0$ is a constant of integration. The boundary conditions on the plate are

$$u_0 = 0 \qquad \text{on } z = 0,\ \zeta_0(y). \tag{53}$$

The final condition on the system is constant imposed volume flux[4]

$$\frac{1}{6}Q = \int_{-1}^{1}\int_{0}^{\zeta_0(y)} u_0(y', z)\,\mathrm{d}z\,\mathrm{d}y', \tag{54}$$

where the flux has been non-dimensionalized by $\rho g \sin\theta\, y_0^4/6\mu$.

Equation (51) subject to the boundary condition (53) is a complex system to solve, more so as the solution for ζ_0 depends on condition (54). However, in view of the scalings (8) used to obtain the von Kármán equations, the vertical length-scale for the fluid must be much less than the horizontal. Hence a lubrication-like approximation may be made and (51) replaced by

$$u_{0,zz} = -1,$$

with solution satisfying (53) given by

$$u = z[\zeta_0(y) - z].$$

Condition (54) thus becomes

$$Q = \int_{-1}^{1} \zeta_0^3(y)\,\mathrm{d}y. \tag{55}$$

We now look at solutions for the governing system of equations (48)–(53) and (55). First we consider the solution for ζ_0. Substituting (52) into (49) and solving for ζ_0 subject to the boundary conditions (50c,d) we obtain

$$\zeta_0(y) = \tilde{p}_0 \quad \frac{-m_2 \sinh m_2 \cosh m_1 y + m_1 \sinh m_1 \cosh m_2 y}{m_2 \sinh m_2 \cosh m_1 - m_1 \sinh m_1 \cosh m_2} + 1 \quad , \tag{56}$$

where

$$m_{1,2} = \sqrt{\frac{\mathfrak{E}_r N_0 \pm \sqrt{\mathfrak{E}_r^2 N_0^2 - \frac{4}{3}\delta_r^2 \mathfrak{E}_r}}{\frac{2}{3}\delta_r^2 \mathfrak{E}_r}}.$$

The behaviour of this solution depends on the sign and magnitude of $\mathfrak{E}_r N_0$. For $\mathfrak{E}_r N_0 < -\sqrt{4\delta_r^2 \mathfrak{E}_r/3}$, m_i are purely imaginary, and the hyperbolic cosine terms in (56) become cosines, giving an oscillating solution for ζ_0. For $\mathfrak{E}_r N_0 > \sqrt{4\delta_r^2 \mathfrak{E}_r/3}$, m_i are real and the solution consists of hyperbolic cosine terms, and lacks oscillations. For $|\mathfrak{E}_r N_0| < \sqrt{4\delta_r^2 \mathfrak{E}_r/3}$,

[4] If the system were on a flat base, then this final condition would be a fixed volume constraint

$$V = \int_{-1}^{1} \zeta_0(y)\,\mathrm{d}y.$$

The analysis for this case is similar to that presented for the fixed flux. It is also related to work by Riera and Mahadevan [23] on the drying of sessile drops.

the solution for ζ_0, (56), becomes a sum of products of hyperbolic cosines and cosines. Note that, regardless of the value of $\mathfrak{E}_r N_0$, the form in (56) is real.[5]

N_0 and $\tilde{p}_0$ are still unknowns that must be specified. If we assume N_0 is known, then $\tilde{p}_0$ may be found from (55) to give

$$\tilde{p}_0 = \quad Q \Big/ \int_{-1}^{1} (\zeta_0/\tilde{p}_0)^3 \, \mathrm{d}y \quad ^{\frac{1}{3}}. \tag{57}$$

This eliminates the singularity in the denominator of ζ_0 in (56) at $m_1 \tanh m_1 = m_2 \tanh m_2$, however it introduces a new one where $\int_{-1}^{1} (\zeta_0/\tilde{p}_0)^3 \, \mathrm{d}y = 0$.

Finally, N_0 may be found by solving (48b) for η_0. Subject to boundary condition (50b), the solution is

$$\eta_0 = \frac{y+1}{4} \int_{-1}^{1} \zeta_{0,y}^2 \, \mathrm{d}y - \frac{1}{2} \int_{-1}^{y} \zeta_{0,y'}^2 \, \mathrm{d}y',$$

from which we obtain

$$N_0 = \epsilon_0 + \frac{1}{4} \int_{-1}^{1} \zeta_{0,y}^2 \, \mathrm{d}y. \tag{58}$$

Hence we may treat the problem as an inverse one: we fix the dependent variable N_0, which provides the full solution for ζ_0, and calculate the imposed parameter ϵ_0 as a function of N_0 using (58).

We may write this relation

$$\mathfrak{E}_r \epsilon_0 = \mathfrak{E}_r N_0 - \frac{(\mathfrak{E}_r^{3/2} Q)^{2/3} \int_{-1}^{1} (\zeta_{0,y}/\tilde{p}_0)^2 \, \mathrm{d}y}{\int_{-1}^{1} (\zeta_{0,y}/\tilde{p}_0)^3 \, \mathrm{d}y \quad ^{2/3}}.$$

Hence note that by scaling the parameters, ϵ_0, Q and δ_r, and the dependent variables, N_0, $\tilde{p}_0$ and ζ_0, appropriately, we eliminate explicit dependence on the parameter $\mathfrak{E}_r$.

Figure 10(a) shows a contour plot of the scaled pre-strain, $\mathfrak{E}_r \epsilon_0$, against $\mathfrak{E}_r N_0$ and the scaled volume flux, $\mathfrak{E}_r^{3/2} Q$. For a given $\mathfrak{E}_r \epsilon_0 < 0$, and a given value of $\mathfrak{E}_r^{3/2} Q$, there are multiple possible values of $\mathfrak{E}_r N_0$. The system has non-unique solutions: a feature of buckling.

Figure 10(b) shows the full set of solutions for the choice of parameters $\mathfrak{E}_r^{3/2} Q = 1$ and $\mathfrak{E}_r \epsilon_0 = -3$. Solutions which go below the y-axis are physically not allowed, however there still remain several valid solutions. The green curve (plotted with $\zeta_0(-1) = 4.5$) is the solution we choose for the system, based on an energy minimization argument, as discussed below.

There are some other notable features of the plot. The solid, black vertical lines satisfy

$$\int_{-1}^{1} (\zeta_0/\tilde{p}_0)^3 \, \mathrm{d}y = 0, \tag{59}$$

[5]There are a couple of other points to note. Firstly, the form of the solution (56) is single-valued although this is not necessarily physically correct. Secondly, if δ_r becomes too small, then the flexural term should be neglected from the plate equations (21). Note also, that if there are too many wiggles, then the assumption that the horizontal length-scale for the plate is much larger than the vertical also fails.

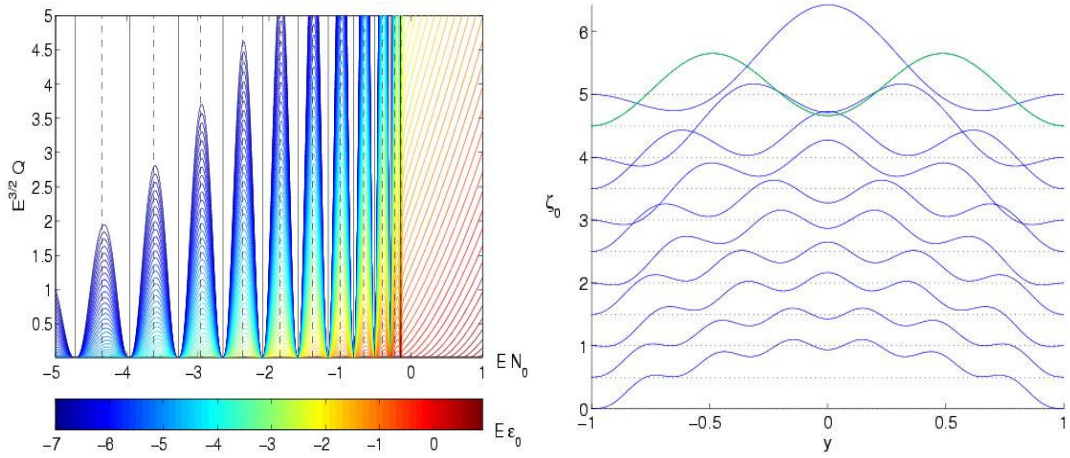


Figure 10: (a) Contours of the imposed parameter $\mathfrak{E}_r\epsilon_0$ against the dependent variable $\mathfrak{E}_r N_0$ and parameter $\mathfrak{E}_r^{3/2}Q$. For a given value of $\mathfrak{E}_r\epsilon_0$ there corresponds a contour. Then, for a given value of $\mathfrak{E}_r^{3/2}Q$ there are multiple possible values of $\mathfrak{E}_r N_0$, and hence multiple possible solutions for ζ_0. The solid and dashed vertical lines satisfy (59) and (60) respectively. (b) Solutions of ζ_0 for the choice of parameters $\mathfrak{E}_r^{3/2}Q = 1$ and $\mathfrak{E}_r\epsilon_0 = -3$. The lowest solution corresponds to the smallest solution for $\mathfrak{E}_r N_0$, and higher up solutions correspond to larger values of $\mathfrak{E}_r N_0$. The green curve (plotted with $\zeta_0(-1) = 4.5$) is the curve chosen on the basis of minimum energy.

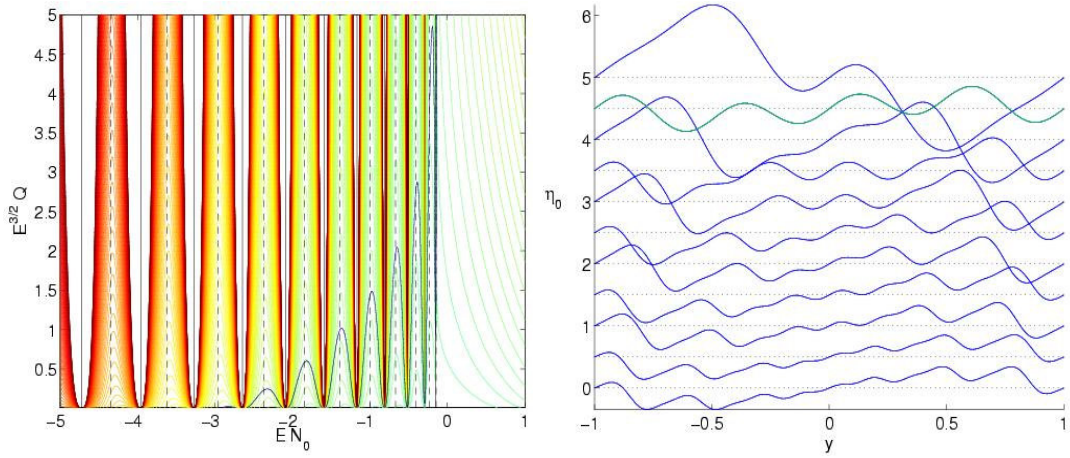


Figure 11: (a) Contours of plot of energy against $\mathfrak{E}_r N_0$ and $\mathfrak{E}_r^{3/2}Q$ together with the contour $\mathfrak{E}_r\epsilon_0 = -3$. The solid and dashed lines again satisfy (59) and (60) respectively. (b) Solutions of η_0 for the choice of parameters $\mathfrak{E}_r^{3/2}Q = 1$ and $\mathfrak{E}_r\epsilon_0 = -3$. The lowest solution corresponds to the smallest solution for $\mathfrak{E}_r N_0$, and higher up solutions correspond to larger values of $\mathfrak{E}_r N_0$. The green curve is the curve chosen on the basis of minimum energy.

or equivalently are where $\tilde{p}_0$ is infinite from (57). Across these lines, the solution ζ_0 is reflected in the y-axis. The dashed black lines satisfy

$$\frac{1}{m_1}\tanh m_1 = \frac{1}{m_2}\tanh m_2 \tag{60}$$

and occur when $\zeta_{0,yy}/\tilde{p}_0 = 0$ at $y = \pm 1$. In regions with solid lines to the left and dashed lines to the right, $\zeta_{0,yy} < 0$ at $y = \pm 1$ and in regions with dashed lines to the left and solid lines to the right, $\zeta_{0,yy} > 0$ at $y = \pm 1$. Hence only the latter are physically acceptable. At dashed lines, as $\mathfrak{E}N_0$ increases, two extra down-wiggles are created. At the next solid line, the shape is reflected so the down-wiggles become up-wiggles. Hence, beginning with one up-wiggle and zero down-wiggles in the right-most finger of figure 10(a), there is one more up-wiggle in each finger to the left.

The choice of solution for a given set of parameters must still be made. We select the minimum energy solution.

The total energy of the system is divided into two parts: the internal (stretching and bending) energy of the elastic plate and the potential energy of the fluid. The two components of the plate energy per unit area are given by (16)

$$\psi_{\text{bending}} = \frac{1}{3}\delta_r^2 \mathfrak{E}_r \zeta_{0,yy}^2, \quad \psi_{\text{stretching}} = \mathfrak{E}_r \;\; \frac{1}{2}(1-\nu)\xi_{0,y}^2 + N_0^2$$

respectively, where the energies have been non-dimensionalized by $\rho g \cos\theta \, y_0^2/2$. The potential energy of the fluid per unit area is

$$\psi_{\text{potential}} = \zeta_0^2.$$

For zero Reynolds' number flows, the kinetic energy is negligible. Hence the total energy per unit length becomes

$$\mathfrak{E}_r \Psi_{\text{total}} = \frac{1}{3}\delta_r^2 \mathfrak{E}_r \int_{-1}^{1} \mathfrak{E}_r \zeta_{0,yy}^2 \, \mathrm{d}y + 2(\mathfrak{E}_r N_0)^2 + \frac{1}{2}(1-\nu)\int_{-1}^{1} \mathfrak{E}_r^2 \xi_{0,y}^2 \, \mathrm{d}y + \int_{-1}^{1} \mathfrak{E}_r \zeta_0^2 \, \mathrm{d}y.$$

Note that this expression again has no explicit dependence on $\mathfrak{E}_r$.

Figure 11(a), shows a plot of energy contours against $\mathfrak{E}_r N_0$ and $\mathfrak{E}_r^{3/2} Q$, together with the contour $\mathfrak{E}_r \epsilon_0 = -3$. In this case, the solution with minimum energy is seen to correspond to the one with least wiggles. This is because there is little compressive stress remaining in this case. This analysis is perhaps not quite correct though: the cross-stream displacement η_0 is shown in figure 11(b) and it is seen to increase as the number of wiggles decreases. It might have been expected that more wiggles would be preferable so that the plate fits more easily over a given volume of fluid.

5.3.2 Linearized perturbation

We now consider the stability of this base state to small perturbations. The linearized perturbation equations for the plate are

$$\frac{2d\mathcal{E}}{1-\nu^2}(\xi_{,xx} + \nu\eta_{,xy} + \nu\zeta_{0,y}\zeta_{,xy}) + \frac{d\mathcal{E}}{1+\nu}(\xi_{,yy} + \eta_{,xy} + \zeta_{0,yy}\zeta_{,x} + \zeta_{0,y}\zeta_{,xy}) = \mu u_{0,zz}\zeta,$$
$$\frac{d\mathcal{E}}{1+\nu}(\xi_{,xy} + \eta_{,yy} + \zeta_{0,y}\zeta_{,xx}) + \frac{2d\mathcal{E}}{1-\nu^2}(\nu\xi_{,xy} + \eta_{,yy} + \zeta_{0,yy}\zeta_{,y} + \zeta_{0,y}\zeta_{,yy}) = 0$$

for the in-plane plate equations (20) and

$$B\nabla_H^4\zeta = p_{0,z}\zeta + 2\mu u_{0,z}\zeta_{,x} + p - 2\mu w_{,z} + \rho g\cos\theta\, y_0^2 N_0(\nu\zeta_{,xx} + \zeta_{,yy}) + 2\frac{d\mathcal{E}}{1+\nu}\xi_{0,y}\zeta_{,xy} + \frac{2d\mathcal{E}}{1-\nu^2}\zeta_{0,yy}(\nu\xi_{,x} + \eta_{,y} + \zeta_{0,y}\zeta_{,y}) + \frac{d\mathcal{E}}{1+\nu}\xi_{0,yy}\zeta_{,x}$$

for the flexural equation (21), with boundary conditions (27) and (28)

$$\xi = \eta = \zeta = \zeta_{,y} = 0 \qquad \text{at } y = \pm y_0.$$

The linearized perturbation equations for the fluid (22) are

$$\mu\nabla^2 u = p_{,x}, \quad \mu\nabla^2 v = p_{,y}, \quad \mu\nabla^2 w = p_{,z}, \quad u_{,x} + v_{,y} + w_{,z} = 0,$$

with boundary conditions

$$u = v = w = 0 \qquad \text{on } z = 0;\ y = \pm y_0.$$

Finally, the linearized interface conditions (23) are

$$0 = u + u_{0,z}\zeta, \quad 0 = v, \quad \zeta_{,t} = w \quad \text{on } z = H.$$

We set $L_p' < \zeta_0(0)$ and $L_p < y_0$ to be the out-of-plane and in-plane length-scales of the perturbation repectively. Then we have neglected terms that are $\mathcal{O}(L_\mathrm{p}'/L_\mathrm{p})$, smaller than the leading order terms in each of these equations, as in §5.2.2.

Assuming perturbations are small wavelength and background variation of the base state is slow, we substitute solutions proportional to $e^{ikx+ily-i\omega t}$ into the above equations and derive a dispersion relation with real part

$$2\mu K\frac{\cosh K\zeta_0 \sinh K\zeta_0 + K\zeta_0}{\sinh^2 K\zeta_0 - K^2\zeta_0^2}\,\Im(\omega) = -BK^4 - \rho g\cos\theta - \rho g\cos\theta\, y_0^2 N_0(\nu k^2 + l^2) - 2\frac{d\mathcal{E}}{1+\nu}\xi_{0,y}kl - \frac{d\mathcal{E}}{K^2(1-\nu^2)}\zeta_{0,yy}^2[(1-\nu)k^2 + 2l^2],$$

where $K^2 = k^2 + l^2$.

This has similar properties to the displacement relation (47) of §5.2.2. The first two terms arise from the flexural and hydrostatic terms respectively and are stabilizing. The third term arises from 'remaining' slack in the system, and can be destabilizing provided the solution for N_0 in the base state is negative. If N_0 is sufficiently large and positive (the system is very taut), then it may eliminate all instability. As in §5.2.2, the base state downstream stretching can be destabilizing and is responsible for symmetry breaking. Much more work must be done to complete this analysis.

6 Conclusions and future work

We have presented some preliminary analysis on an elastic-skinned gravity current and the wrinkling phenomenon that may be observed, under certain conditions, on the surface. This

wrinkling is most prominent due to a combination of a compressive pre-strain on the plate and fluid traction pulling at the base of the plate. The pulling is what is responsible for symmetry breaking. We have tried to remain consistent with the scalings under which the von Kármán plate equations hold. Some experimental observations have been explained from our analysis: the necessity of slack for wrinkles to be observed and the downstream orientation of the wrinkles.

There is still very much that can be done. The stability analyses can be completed in a WKB sense and can be considered numerically. Beyond that, quantitative experiments and numerical simulations using finite elements could be performed. Formulating boundary conditions for real elastic-skinned gravity currents should also be done.

7 Acknowledgements

I am extremely grateful to Neil Balmforth for all his time and patient help and Andrew Belmonte for showing me how much fun it is to be an experimentalist. A huge thank you to all the other Fellows and pseudo-Fellows for their food, tea and especially their company that made for a great summer. Finally, thank you to everyone at the Cottage for making it such an exciting place.

References

[1] M. Clayton, Master's thesis, Penn State, 2004.

[2] C. D. Ollier, *Volcanoes* (MIT Press, Cambridge, MA, 1969).

[3] J. H. Fink and R. C. Fletcher, "Ropy pahoehoe: surface folding of a viscous liquid," J. Volcanol. Geoth. Res. **4**, 151 (1978).

[4] H. E. Huppert, "Flow and instability of a viscous gravity current down a slope," Nature **300**, 427 (1982).

[5] J. B. Grotberg and O. E. Jensen, "Biofluid mechanics in flexible tubes," Annu. Rev. Fluid Mech. **36**, 121 (2004).

[6] D. J. Gee, "Numerical continuation applied to panel flutter," Nonlinear Dynamics **22**, 271 (2000).

[7] T. B. Benjamin, "Effects of a flexible boundary on hydrodynamic stability," J. Fluid Mech. **9**, 513 (1960).

[8] A. Chakrabarti, D. S. Ahluwalia, and S. R. Manam, "Surface water waves involving a vertical barrier in the presence of an ice-cover," Int. J. Eng. Sci. **41**, 1145 (2003).

[9] N. J. Balmforth, R. V. Craster, and A. C. Rust, "Instability in flow through elastic conduits and volcanic tremor," J. Fluid Mech. (2004), submitted.

[10] A. E. Hosoi and L. Mahadevan, "Peeling, healing and bursting in a lubricated elastic sheet," Phys. Rev. Lett. (2004), to appear.

[11] Y.-Y. Yu, *Vibrations of elastic plates: linear and non-linear dynamical modelling of sandwiches, laminated composites, and piezoelectric layers* (Springer, New York, 1996).

[12] Y. C. Fung, *Foundations of solid mechanics* (Prentice-Hall, New Delhi, 1968).

[13] S. P. Timoshenko and S. Woinowsky-Krieger, *Theory of plates and shells* (McGraw-Hill, New York, 1959).

[14] L. D. Landau and E. M. Lifschitz, *Theory of elasticity* (Pergamon, Oxford, 1970).

[15] P. G. Ciarlet, "Une justification des équations de von Kármán," C. R. Acad. Sc. Paris Série A **288**, 469 (1979).

[16] O. Millet, A. Hamdouni, and A. Cimetière, "A classification of thin plate models by asymptotic expansion of non-linear three-dimensional equilibrium equations," Non-linear Mechanics **36**, 165 (2001).

[17] S. S. Antman, *Nonlinear problems of elasticity* (Springer, New York, 1995).

[18] J. J. Stoker, *Nonlinear elasticity* (Nelson, London, 1968).

[19] R. Huang and Z. Suo, "Wrinkling of a compressed elastic film on a viscous layer," J. Appl. Phys. **91**, 1135 (2002).

[20] Y. W. Wong and S. Pellegrino, in *43rd AIAA/ASME/ASCE/AHS/ASC Structures, structural dynamics, and materials conference* (AIAA, Denver, CO, 2002).

[21] E. H. Mansfield, *The bending and stretching of plates*, *International series of monographs on aeronautics and astronautics* (Pergamon, Oxford, 1964).

[22] Y. W. Wong and S. Pellegrino, in *New approaches to structural mechanics, shells and biological structures*, edited by H. Drew and S. Pellegrino (Kluwer, Dordrecht, 2002), Chap. Amplitude of wrinkles in thin membrane.

[23] C. S. Riera, `http://people.deas.harvard.edu/~riera/drying.html`.

High-order Boussinesq models for internal interfacial waves

Yaron Toledo

Technion - Israel Institute of Technology

1 Introduction

The irrotational flow of an incompressible homogenous inviscid fluid is generally a three-dimensional problem. The main issue of Boussinesq-type equations is to reduce the description to a two-dimensional one by introducing a polynomial approximation to the vertical distribution of the flow into integral conservation laws, while accounting for non-hydrostatic effects due to the vertical acceleration of water. After solving the two-dimensional problem we can easily find the flow properties everywhere in the three-dimensional domain using the polynomial approximation. In this work we apply this method to the flow of internal interfacial waves between two incompressible fluids. In the first part of the work (sections 2 and 3) both fluids are homogeneous and the flow is irrotational either with a rigid-lid or a free surface. In the second part (section 4) the fluids are both exponentially stratified and the flow is irrotational with a rigid-lid (a similar problem with a free surface could be solved as well by using the same techniques). In the third part of the work (section 5) the method is applied to the rotational flow, which is presented by the surface quasi-geostrophic model. Finally, in the last part of the work (section 6) a new layered model based on the surface quasi-geostrophic model is constructed and its dimension is reduced by the Boussinesq method.

2 Internal interfacial waves between two unstratified layers with a rigid-lid

2.1 An infinite series solution for the Laplace equation in each layer over a horizontal bottom

The equations governing the irrotational flow of an incompressible inviscid fluid in the lower layer over a horizontal bottom are:

$$\nabla^2\Phi^{(1)} + \Phi^{(1)}_{zz} = 0 \qquad -h_1 < z < \eta \tag{1}$$

$$\eta_t + \nabla\Phi^{(1)}\nabla\eta - \Phi^{(1)}_z = 0 \qquad z = \eta \tag{2}$$

$$\Phi^{(1)}_t + \frac{1}{2}\left(\nabla\Phi^{(1)}\right)^2 + \frac{1}{2}\left(\Phi^{(1)}_z\right)^2 + g\eta - \frac{P}{\rho_1} = 0 \qquad z = \eta \tag{3}$$

$$\Phi^{(1)}_z = 0 \qquad z = -h_1 \tag{4}$$

The equations governing the irrotational flow of an incompressible inviscid fluid in the upper layer under a rigid lid are:

$$\nabla^2\Phi^{(2)} + \Phi^{(2)}_{zz} = 0 \qquad \eta < z < h_2 \tag{5}$$

$$\eta_t + \nabla\Phi^{(2)}\nabla\eta - \Phi^{(2)}_z = 0 \qquad z = \eta \tag{6}$$

$$\Phi^{(2)}_t + \frac{1}{2}\left(\nabla\Phi^{(2)}\right)^2 + \frac{1}{2}\left(\Phi^{(2)}_z\right)^2 + g\eta - \frac{P}{\rho_2} = 0 \qquad z = \eta \tag{7}$$

$$\Phi^{(2)}_z = 0 \qquad z = h_2 \tag{8}$$

By the use of (2), (3), (6), and (7) we can construct two interface conditions on $z = \eta$:

$$\rho_1\left(\Phi^{(1)}_t + \tfrac{1}{2}\left(\nabla\Phi^{(1)}\right)^2 + \tfrac{1}{2}\left(\Phi^{(1)}_z\right)^2 + g\eta\right) = \rho_2\left(\Phi^{(2)}_t + \tfrac{1}{2}\left(\nabla\Phi^{(2)}\right)^2 + \tfrac{1}{2}\left(\Phi^{(2)}_z\right)^2 + g\eta\right) \tag{9}$$

$$-\nabla\Phi^{(1)}\nabla\eta + \Phi^{(1)}_z = -\nabla\Phi^{(2)}\nabla\eta + \Phi^{(2)}_z \tag{10}$$

Here Φ is the velocity potential, $h_1$and h_2 are the water depths of the two layers, P is the pressure and η the interface elevation. The horizontal gradient operator relates Φ to the horizontal velocity, u:

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right), \quad \mathbf{u} = (u, v). \tag{11}$$

For convenience we denote $\Phi_z = W$ and use the hat ($\hat{}$) and tilde ($\tilde{}$) signs to denote the value on $z = 0$ and on $z = \eta$ respectively.

One of the main ideas of Boussinesq-type theories is to reduce the three-dimensional description to a two-dimensional one. The first step towards such a reduction is to introduce an expansion of the velocity potential as a power of series in the vertical coordinates:

$$\Phi(x, y, z, t) = \sum_{n=0}^{\infty} z^n \phi_n(x, y, t). \tag{12}$$

By substituting this expansion into (1) we find

$$\Phi^{(1)}(x, y, z, t) = \sum_{n=0}^{\infty}(-1)^n \left(\frac{z^{2n}}{(2n)!}\nabla^{2n}\phi^{(1)}_0 + \frac{z^{2n+1}}{(2n+1)!}\nabla^{2n}\phi^{(1)}_1\right). \tag{13}$$

This is a series solution with two unknown functions $\phi^{(1)}_0$and $\phi^{(1)}_1$. Note that the velocities in layer 1 at the undisturbed interface are given by

$$\hat{\mathbf{u}}^{(1)} = \nabla\hat{\Phi}^{(1)} = \nabla\phi_0^{(1)}, \quad \hat{W}^{(1)} = \phi_1^{(1)} \tag{14}$$

Now by the use of (13) and (14) the horizontal bottom condition (4) can be expressed as

$$L_c\left\{\hat{W}^{(1)}\right\} + L_s \cdot \left\{\nabla\hat{\Phi}^{(1)}\right\} = 0 \tag{15}$$

with

$$L_c = \sum_{n=0}^{\infty}(-1)^n \frac{h_1^{2n}}{(2n)!}\nabla^{2n}, \qquad L_s = \sum_{n=0}^{\infty}(-1)^n \frac{h_1^{2n+1}}{(2n+1)!}\nabla^{2n+1} \tag{16}$$

where ∇ is the gradient operator when applied to a scalar, and the divergence when applied to a vector. This equation defines a relation between $\hat{W}^{(1)}$ and $\hat{\Phi}^{(1)}$ which is of infinite order in $h\nabla$. The series are convergent if $\Phi^{(1)}$ has a Fourier transform, since they correspond to the analytic functions $\sinh(kh)$ and $\cosh(kh)$ where ik is the Fourier symbol of ∇.

Following Rayleigh (1876) , we may use symbolic notation of Taylor series operators by which (16) can be given in the compact form

$$L_c = \cos(h_1\nabla), \qquad L_s = \sin(h_1\nabla) \tag{17}$$

so that (15) becomes

$$\cos(h_1\nabla)\hat{W}^{(1)} + \sin(h_1\nabla)\nabla\hat{\Phi}^{(1)} = 0 \tag{18}$$

and (13) and its z derivative become

$$\Phi^{(1)}(x,y,z,t) = \cos(z\nabla)\hat{\Phi}^{(1)} + \tfrac{\sin(z\nabla)}{\nabla}\hat{W}^{(1)}, \tag{19}$$

$$W^{(1)}(x,y,z,t) = -\sin(z\nabla)\nabla\hat{\Phi}^{(1)} + \cos(z\nabla)\hat{W}^{(1)}. \tag{20}$$

Using (18) we can easily construct a Dirichlet to Neumann relation

$$\hat{W}^{(1)} = -\tan(h_1\nabla)\nabla\hat{\Phi}^{(1)} \quad @z = 0 \tag{21}$$

and define its operator as

$$G_1 = -\tan(h_1\nabla)\nabla. \tag{22}$$

By applying the same techniques we can construct a similar Dirichlet to Neumann relation for the second layer

$$\hat{W}^{(2)} = \tan(h_2\nabla)\nabla\hat{\Phi}^{(2)} \quad @z = 0 \tag{23}$$

2.2 Constructing the accurate equations for the linear problem

For small amplitude waves in the sense that the ration between the amplitude to the wave length (wave slope) is small the problem becomes linear. We can see that the kinematic and dynamic interface conditions (10) and (9) become

$$\hat{W}^{(1)} = \hat{W}^{(2)} \quad @z = 0 \tag{24}$$

$$\rho_1 \quad \hat{\Phi}_t^{(1)} + g\eta \quad = \rho_2 \quad \hat{\Phi}_t^{(2)} + g\eta \qquad @z = 0 \tag{25}$$

By the use of (21), (23) and (24) we can construct a relation between $\hat{\Phi}^{(1)}$ and $\hat{\Phi}^{(2)}$

$$\hat{\Phi}^{(2)} = -\frac{\tan(h_1\nabla)}{\tan(h_2\nabla)}\hat{\Phi}^{(1)} \quad @z = 0. \tag{26}$$

Let us define this operator as

$$G_2 = -\frac{\tan(h_1\nabla)}{\tan(h_2\nabla)}. \tag{27}$$

By using the linear version of (2), (26), (27) and the derivative of (25) with respect to time we get

$$\left(\rho_1 - \rho_2 G_2\right)\hat{\Phi}_{tt}^{(1)} = \left(\rho_2 - \rho_1\right) g\hat{W}^{(1)} \quad @z = 0 \tag{28}$$

By the use of (21) and (22) equation (28) becomes

$$\left(\rho_1 - \rho_2 G_2\right)\hat{\Phi}_{tt}^{(1)} = \left(\rho_2 - \rho_1\right) g G_1\hat{\Phi}^{(1)} \quad @z = 0 \tag{29}$$

By substituting the linear relations $\nabla = ik$ and $\frac{\partial}{\partial t} = i\omega$ and the definitions of the differential operators (22) and (27) into (29) we get

$$\omega^2 = \frac{\left(\rho_1 - \rho_2\right) g \tanh\left(kh_1\right)\tanh\left(kh_2\right)}{\rho_1 \tanh\left(kh_2\right) + \rho_2 \tanh\left(kh_1\right)} \tag{30}$$

We can see that equation (30) is exactly the well known linear dispersion relation for this problem.

2.3 Constructing approximate equations for the linear problem

One way to use the above equations for solving a general wave problem is by using Fourier and Laplace transforms with the exact linear dispersion (30). A solution using that method have two downsides. The first is the difficulty of computing its integrals and the other downside is that this type of linear solution doesn't help us construct the solution of the nonlinear problem.

We introduce here another method for solving the linear problem. In this method we first need to approximate the infinite differential operators (22) and (27) to finite operators. The accuracy of this solution will relate to the accuracy of the approximations that are used. The simplest way of doing that is by truncating the Taylor series representing these operators. The higher the order of derivatives kept in the Taylor expansions the higher the accuracy. But, there is a better way of approximating these operators.

Padé approximation approximates these operators by a ratio of two power series. It has double the accuracy of the Taylor approximation, while using the same order of derivatives. For clarity we give both the Taylor and the Padé approximations. Taylor (2) represents the Taylor approximation up to the order of ∇^2 (including) and Padé (2,2) represents the Padé approximation up to the order of ∇^2 both in the numerator and the denominator.

Note that the two operators being approximated here are even operators. Therefore, for the Taylor approximation the lowest order term neglected is actually of $O\ \ \nabla^4\ \ $ and for the Padé approximation it is of $O\ \ \nabla^8\ \ $.

The Taylor and Padé approximations for relation (21) are

$$\hat{W}^{(1)} = -h_1 \nabla^2 \hat{\Phi}^{(1)} \qquad \text{Taylor (2)} \tag{31}$$

$$1 - \tfrac{1}{3} h_1^2 \nabla^2\ \ \hat{W}^{(1)} = -h_1 \nabla^2 \hat{\Phi}^{(1)} \quad \text{Padé (2,2).} \tag{32}$$

The Taylor and Padé approximations for the coupling relation (26) are

$$\hat{\Phi}^{(2)} = -\tfrac{h_1}{h_2}\ \ 1 + \tfrac{1}{3}\ \ h_1^2 - h_2^2\ \ \nabla^2\ \ \hat{\Phi}^{(1)} \qquad \text{Taylor (2)} \tag{33}$$

$$15 - \ \ 6h_1^2 + h_2^2\ \ \nabla^2\ \ \hat{\Phi}^{(2)} = -\tfrac{h_1}{h_2}\ \ 15 - \ \ h_1^2 + 6h_2^2\ \ \nabla^2\ \ \hat{\Phi}^{(1)} \quad \text{Padé (2,2)} \tag{34}$$

Now we can use the Taylor (2) approximations to construct an approximated version of equation (29) by using the (31), (33) and the derivative of (25) with respect to time

$$(\rho_1 h_2 + \rho_2 h_1) + \tfrac{1}{3}\rho_2 h_1\ \ h_1^2 - h_2^2\ \ \nabla^2\ \ \hat{\Phi}_{tt}^{(1)} = (\rho_1 - \rho_2)\, h_1 h_2 g \nabla^2 \hat{\Phi}^{(1)} \tag{35}$$

and by using the Padé (2,2) approximations (32) and (34) we can construct in the same manner the following equation

$$\rho_1 h_2 - \rho_2 h_1\ \ 15 - \ \ h_1^2 + 6h_2^2\ \ \nabla^2 \quad \hat{\Phi}_{tt}^{(1)} = (\rho_2 - \rho_1)\, g h_2\ \ 15 - \ \ 6h_1^2 + h_2^2\ \ \nabla^2\ \ \hat{W}^{(1)}. \tag{36}$$

Notice that we are not substituting equation (32) into (36) because by doing so the order of differentiation will increase resulting in a more complicated method. Therefore, we should first use (32) for computing $\hat{W}^{(1)}$ and then only then substitute $\hat{W}^{(1)}$ into (36) and solve for $\hat{\Phi}^{(1)}$.

2.4 Constructing approximate equations for the nonlinear Dirichlet to Neumann problem

From the Laplace equation (1) we can replace horizontal by vertical differentiation to obtain for layer 1

$$\nabla^{2m}\Phi^{(1)} = (-1)^m \frac{\partial^{2m}\Phi^{(1)}}{\partial z^{2m}} \tag{37}$$

Now we look for $\Phi^{(1)}$ of the form

$$\Phi^{(1)}(x,y,z,t) = \sum_{n=0}^{\infty}(z+h_1)^n \phi_n^{(1)}(x,y,t). \tag{38}$$

From (38) it follows that

$$\frac{\partial^m \Phi^{(1)}}{\partial z^m} = \sum_{n=0}^{\infty} \frac{n!}{(n-m)!}(z+h_1)^{n-m}\phi_n^{(1)}(x,y,t) \tag{39}$$

By substituting (38) into the Laplace equation (1) and collecting equal powers of z we obtain a recursive relation

$$\phi_{n+2}^{(1)} = -\frac{\nabla^2 \phi_n^{(1)}}{(n+1)(n+2)}. \tag{40}$$

On the horizontal bottom, boundary condition (4) leads to

$$\phi_1^{(1)} = 0. \tag{41}$$

This implies from (40) that all $\phi_n^{(1)}$'s with odd n vanish

$$\phi_1^{(1)} = \phi_3^{(1)} = \phi_5^{(1)} = ... = 0. \tag{42}$$

Now by the use of (38), (37) and (39) we can write $\hat{\Phi}^{(1)}$, $\nabla^2\hat{\Phi}^{(1)}$, $\hat{W}^{(1)}$ and $\nabla^2\hat{W}^{(1)}$, which relate to the properties of the flow on $z = 0$ as

$$\begin{aligned} \hat{\Phi}^{(1)} &= \sum_{n=0}^{3} h_1^{2n}\phi_{2n}^{(1)} \qquad & \hat{W}^{(1)} &= \sum_{n=0}^{3} 2n h_1^{2n-1}\phi_{2n}^{(1)} \\ \nabla^2\hat{\Phi}^{(1)} &= -\sum_{n=0}^{3} \frac{(2n)!}{(2n-2)!} h_1^{2n}\phi_{2n}^{(1)} \qquad & \nabla^2\hat{W}^{(1)} &= -\sum_{n=0}^{3} \frac{(2n)!}{(2n-3)!} h_1^{2n}\phi_{2n}^{(1)} \end{aligned} \tag{43}$$

Notice that (38) was truncated to use only 4 base functions. This enables us to solve for them using the 4 equations written in (43). Had we wanted to use more base functions for higher accuracy, we would have needed to increase the number of equations by adding equations for higher order of differentiation to equation set (43). By solving equation set (43) for the base functions $\phi_{2n}^{(1)}$, $n = 0..3$ in terms of $\hat{\Phi}^{(1)}$ and $\hat{W}^{(1)}$ and their second horizontal derivatives we get

$$
\begin{aligned}
\phi_0^{(1)} &= \frac{1}{48} \quad -9h_1^2\nabla^2\hat{\Phi}^{(1)} + h_1^3\nabla^2\hat{W}^{(1)} - 33h_1\hat{W}^{(1)} + 48\hat{\Phi}^{(1)} \\
\phi_2^{(1)} &= \frac{1}{16} \quad 7\nabla^2\hat{\Phi}^{(1)} - h_1\nabla^2\hat{W}^{(1)} + \frac{15}{h_1}\hat{W}^{(1)} \\
\phi_4^{(1)} &= \frac{1}{16} \quad -\frac{5}{h_1^2}\nabla^2\hat{\Phi}^{(1)} + \frac{1}{h_1}\nabla^2\hat{W}^{(1)} - \frac{5}{h_1^3}\hat{W}^{(1)} \\
\phi_6^{(1)} &= \frac{1}{48} \quad \frac{3}{h_1^4}\nabla^2\hat{\Phi}^{(1)} - \frac{1}{h_1^3}\nabla^2\hat{W}^{(1)} + \frac{3}{h_1^5}\hat{W}^{(1)}
\end{aligned} \tag{44}
$$

Using the solution (44) in (38) on $z = \eta$ we get a relation involving $\hat{\Phi}^{(1)}$, $\hat{W}^{(1)}$ and $\tilde{\Phi}^{(1)}$:

$$
\begin{aligned}
\tilde{\Phi}^{(1)} = \frac{\eta^3 \ \ \eta + 2h_1^2 \ ^3}{48h_1^3}\nabla^2\hat{W}^{(1)} + \frac{\eta^2 \ \ \eta + 2h_1^2 \ ^2 \ \ \eta^2 + 2\eta h_1 - 2h_1^2}{16h_1^4}\nabla^2\hat{\Phi}^{(1)} \\
+\frac{16h_1^5\eta + 10h_1^2\eta^4 + 6h_1\eta^5 + \eta^6}{16h_1^5}\hat{W}^{(1)} + \hat{\Phi}^{(1)}
\end{aligned} \tag{45}
$$

Equations (15) and (45) are two equations for $\hat{\Phi}^{(1)}$ and $\hat{W}^{(1)}$. We can solve for $\hat{\Phi}^{(1)}$ and $\hat{W}^{(1)}$using a numerical method such as finite differences. The set of equation will have the form

$$
\begin{bmatrix} A_{\text{padé}}^{(1)} & -B_{\text{padé}}^{(1)} \\ A_{\text{nl}}^{(1)} & B_{\text{nl}}^{(1)} \end{bmatrix} \begin{matrix} \hat{\Phi}^{(1)} \\ \hat{W}^{(1)} \end{matrix} = \begin{matrix} \mathbf{0} \\ \tilde{\Phi}^{(1)} \end{matrix} \quad . \tag{46}
$$

Here $A_{\text{padé}}^{(1)}$ and $B_{\text{padé}}^{(1)}$ are the finite difference matrices representing the Dirichlet to Neumann relation (32), $A_{\text{nl}}^{(1)}$ and $B_{\text{nl}}^{(1)}$ are the finite difference matrices representing the relation between the undisturbed potential and vertical velocity and the potential on the interfacial wave. Now by using $\hat{\Phi}^{(1)}$ and $\hat{W}^{(1)}$ we can find the base functions and from the base functions we can find the properties of the flow at every point in the layer.

Now let us present a way to calculate the properties of the flow in the upper layer. We can easily find $\tilde{W}^{(2)}$ using (10). Applying the same technique described in subsection 2.1 we can construct an equivalent version of equations (19) and (20) for the upper layer

$$
\Phi^{(2)}(x,y,z,t) = \cos(z\nabla)\hat{\Phi}^{(2)} + \tfrac{\sin(z\nabla)}{\nabla}\hat{W}^{(2)}, \tag{47}
$$

$$
W^{(2)}(x,y,z,t) = -\sin(z\nabla)\nabla\hat{\Phi}^{(2)} + \cos(z\nabla)\hat{W}^{(2)}. \tag{48}
$$

On $z = \eta$we can use equations (47) and (48) to solve $\hat{\Phi}^{(2)}$ and $\hat{W}^{(2)}$ and substitute the solution into (47) and (48) to get $\Phi^{(2)}$ and $W^{(2)}$ at every point in the layer.

2.5 Constructing the equations for the nonlinear time marching

An analytical relation between $\tilde{\Phi}_t^{(1)}$ and $\tilde{\Phi}_t^{(2)}$must be developed in order to enable marching $\tilde{\Phi}$ in time using the dynamical interface boundary condition (9). By applying the solution

(44) to the z derivative of (38) on $z=\eta$ we get a second relation between $\hat{\Phi}^{(1)}$ and $\hat{W}^{(1)}$with respect to $\tilde{W}^{(1)}$. By using this relation, equation (45) and equation (15) we can eliminate $\hat{\Phi}^{(1)}$ and $\hat{W}^{(1)}$and obtain an analytical Dirichlet to Neumann relation on the interfacial wave

$$1-\quad \frac{1}{2}\eta^2+\eta h_1 \quad \nabla^2 \quad \tilde{\Phi}^{(1)} = -\left(\eta+h_1\right)\nabla^2\tilde{W}^{(1)} \qquad \text{Taylor (2)} \tag{49}$$

$$1-\quad \frac{1}{2}\eta^2+\eta h_1+\frac{3}{7}h_1^2 \quad \nabla^2 \quad \tilde{\Phi}^{(1)} = -\left(\eta+h_1\right)\nabla^2\tilde{W}^{(1)} \qquad \text{Padé (2,2)} \tag{50}$$

A similar relation can be constructed for the second layer

$$1-\quad \frac{1}{2}\eta^2-\eta h_2 \quad \nabla^2 \quad \tilde{\Phi}^{(2)} = -\left(\eta-h_2\right)\nabla^2\tilde{W}^{(2)} \qquad \text{Taylor (2)} \tag{51}$$

$$1-\quad \frac{1}{2}\eta^2-\eta h_2+\frac{3}{7}h_2^2 \quad \nabla^2 \quad \tilde{\Phi}^{(2)} = -\left(\eta-h_2\right)\nabla^2\tilde{W}^{(2)} \qquad \text{Padé (2,2)} \tag{52}$$

Now by substituting equations (49) and (51) into the kinematic interface boundary condition (10) we get a relation between $\tilde{\Phi}^{(1)}$ and $\tilde{\Phi}^{(2)}$ for Taylor (2) approximation

$$\begin{aligned}\left(\eta-h_2\right) \quad 1-\nabla^2\eta\nabla- \quad \frac{1}{2}\eta^2+\eta h_1+2\nabla^2\eta \quad \nabla^2 \quad \tilde{\Phi}^{(1)} \\ =\left(\eta+h_1\right) \quad 1-\nabla^2\eta\nabla- \quad \frac{1}{2}\eta^2-\eta h_2+2\nabla^2\eta \quad \nabla^2 \quad \tilde{\Phi}^{(2)}\end{aligned} \tag{53}$$

deriving this relation with respect to time yields a relation between $\tilde{\Phi}_t^{(1)}$ and $\tilde{\Phi}_t^{(2)}$
For Taylor (2) approximation

$$\begin{aligned}&\left(\eta-h_2\right) \;\; 1-\nabla^2\eta\nabla- \;\; \tfrac{1}{2}\eta^2+\eta h_1+2\nabla^2\eta \;\; \nabla^2 \;\; \tilde{\Phi}_t^{(1)}+ \\ &\qquad -\eta_t+ \;\; \nabla^3\eta_t\left(\eta-h_2\right)+\nabla^3\eta\eta_t \;\; \nabla \;\; \tilde{\Phi}^{(1)} \\ &+ \;\; 2\nabla^3\eta_t\left(\eta-h_2\right)+ \;\; 2\nabla^2\eta+\tfrac{3}{2}\eta^2-h_2\eta-h_1\left(-2\eta+h_2\right) \;\; \eta_t \;\; \nabla^2\tilde{\Phi}^{(1)} \\ &=\left(\eta+h_1\right) \;\; 1-\nabla^2\eta\nabla- \;\; \tfrac{1}{2}\eta^2-\eta h_2+2\nabla^2\eta \;\; \nabla^2 \;\; \tilde{\Phi}_t^{(2)}+ \\ &\qquad -\eta_t+ \;\; \nabla^3\eta_t\left(\eta+h_1\right)+\nabla^3\eta\eta_t \;\; \nabla \;\; \tilde{\Phi}^{(2)}+ \\ &2\nabla^3\eta_t\left(\eta+h_1\right)+ \;\; 2\nabla^2\eta+\tfrac{3}{2}\eta^2+h_1\eta-h_2\left(2\eta+h_1\right) \;\; \eta_t \;\; \nabla^2\tilde{\Phi}^{(2)}\end{aligned} \tag{54}$$

which together with the dynamical interface boundary condition (9) enable us to solve for $\tilde{\Phi}_t^{(1)}$ and $\tilde{\Phi}_t^{(2)}$ and then propagate them in time. The above derivatives should be handle with care because $\tilde{\Phi}^{(1)}$ and $\tilde{\Phi}^{(2)}$ are located on the interfacial wave, which means that in each location they are located on a different elevation (η). The horizontal derivatives can be locally related to the tangential ($\nabla\tilde{\Phi}\cdot\hat{t}$) derivatives, the elevation (η) and the vertical derivatives $\hat{W}^{(1)}$ and $\hat{W}^{(2)}$ using this relation

$$\nabla\tilde{\Phi}=\left(\nabla\tilde{\Phi}\cdot\hat{t}\right)+ \quad \tilde{W}-\frac{\nabla\eta}{\left|\langle 1,\nabla\eta\rangle\right|} \quad \nabla\tilde{\Phi}\cdot\hat{t} \quad \frac{\nabla\eta}{\left|\langle 1,\nabla\eta\rangle\right|} \tag{55}$$

3 Internal interfacial waves between two unstratified layers with a free surface

3.1 An infinite series expansion for Φ in each layer over a horizontal bottom

The equations governing the irrotational flow of an incompressible inviscid fluid in the lower layer over a horizontal bottom are shown in equations (1)-(4). The equations governing the irrotational flow of an incompressible inviscid fluid in the second layer under a free surface boundary condition are:

$$\nabla^2\Phi^{(2)} + \Phi^{(2)}_{zz} = 0 \qquad \eta < z < h_2 \tag{56}$$

$$\eta_t + \nabla\Phi^{(2)}\nabla\eta - \Phi^{(2)}_z = 0 \qquad z = \eta \tag{57}$$

$$\Phi^{(2)}_t + \frac{1}{2}\ \nabla\Phi^{(2)}\ ^2 + \frac{1}{2}\ \Phi^{(2)}_z\ ^2 + g\eta - \frac{P}{\rho_2} = 0 \qquad z = \eta \tag{58}$$

$$\xi_t + \nabla\Phi^{(2)}\nabla\eta - \Phi^{(2)}_z = 0 \qquad z = h_2 + \xi \tag{59}$$

$$\Phi^{(2)}_t + \frac{1}{2}\ \nabla\Phi^{(2)}\ ^2 + \frac{1}{2}\ \Phi^{(2)}_z\ ^2 + g\xi = 0 \qquad z = h_2 + \xi \tag{60}$$

By the use of (2), (3), (57), and (58) we can construct two interface boundary conditions on $z = \eta$:

$$\rho_1\ \ \Phi^{(1)}_t + \tfrac{1}{2}\ \nabla\Phi^{(1)}\ ^2 + \tfrac{1}{2}\ \ \Phi^{(1)}_z\ ^2 + g\eta\ \ = \rho_2\ \ \Phi^{(2)}_t + \tfrac{1}{2}\ \nabla\Phi^{(2)}\ ^2 + \tfrac{1}{2}\ \ \Phi^{(2)}_z\ ^2 + g\eta \tag{61}$$

$$-\nabla\Phi^{(1)}\nabla\eta + \Phi^{(1)}_z = -\nabla\Phi^{(2)}\nabla\eta + \Phi^{(2)}_z \tag{62}$$

Here Φ is the velocity potential, $h_1$and h_2 are the water depths of the two layers, P is the pressure, η is the interface elevation and ξ is the free surface elevation. The horizontal gradient operator relates the horizontal velocity $\mathbf{u}$ to Φ:

$$\nabla = \ \ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}\ \ , \quad \mathbf{u} = (u, v) = \nabla\Phi \tag{63}$$

For convenience we denote

$$\begin{aligned} \hat{\Phi}^{(2)} &= \Phi^{(2)}(z = h_1)\,, \quad \tilde{\Phi}^{(2)} = \Phi^{(2)}(z = h_2 + \xi) \\ \hat{W}^{(2)} &= W^{(2)}(z = h_1)\,, \quad \tilde{W}^{(2)} = W^{(2)}(z = h_2 + \xi) \end{aligned} \tag{64}$$

The work on the Dirichlet to Neumann relation that has been done for the first layer in the previous section holds for this section. Its linear part can be summarize in equations (15) and (17). For the second layer we again introduce an expansion of the velocity potential as a power of series in z:

$$\Phi^{(2)}(x, y, z, t) = \sum_{n=0}^{\infty} \bar{z}^n \phi^{(2)}_n(x, y, t) \tag{65}$$

Here $\bar{z} = z - h_2$. Now by substituting this expansion into (56) we find

$$\Phi^{(2)} = \cos(\bar{z}\nabla)\hat{\hat{\Phi}}^{(2)} + \frac{\sin(\bar{z}\nabla)}{\nabla}\hat{\hat{W}}^{(2)} \tag{66}$$

By using the above equation we can construct a linear relation between the interfacial wave and the free surface

$$\hat{W}^{(2)} = \sin(h_2\nabla)\nabla\hat{\hat{\Phi}}^{(2)} + \cos(h_2\nabla)\hat{\hat{W}}^{(2)} \tag{67}$$

3.2 Constructing accurate equations for the linear problem

For the linear problem that the kinematic and dynamic interface boundary conditions (62) and (61) become

$$\hat{W}^{(1)} = \hat{W}^{(2)} \quad @z = 0, \tag{68}$$

$$\rho_1 \quad \hat{\Phi}_t^{(1)} + g\eta \quad = \rho_2 \quad \hat{\Phi}_t^{(2)} + g\eta \qquad @z = 0. \tag{69}$$

The free surface boundary conditions (59) and (60) become

$$\xi_t - \Phi_z^{(2)} = 0 \qquad @z = h_2, \tag{70}$$

$$\Phi_t^{(2)} + g\xi = 0 \qquad @z = h_2. \tag{71}$$

By substituting (70) into the time derivative of (71) we get

$$\hat{\hat{\Phi}}_{tt}^{(2)} = -g\hat{\hat{W}}^{(2)} \quad @z = h_2. \tag{72}$$

Substituting (68) into (67) yields

$$\hat{\hat{W}}^{(2)} = -\tan(h_2\nabla)\nabla\hat{\hat{\Phi}}^{(2)} + \frac{1}{\cos(h_2\nabla)}\hat{W}^{(1)} \tag{73}$$

By substituting (21) into (73) and then into (66) on $z = 0$ and deriving twice with respect to time we get

$$\hat{\Phi}_{tt}^{(2)} = \frac{1}{\cos(h_2\nabla)}\hat{\hat{\Phi}}_{tt}^{(2)} + \tan(h_1\nabla)\tan(h_2\nabla)\hat{\hat{\Phi}}_{tt}^{(1)} \tag{74}$$

By using the linear version of (57), (72), (74) and the derivative of (69) with respect to time we get

$$\left(\rho_1 - \rho_2\tan(h_1\nabla)\tan(h_2\nabla)\right)\hat{\Phi}_{tt}^{(1)} = -\rho_2\cos^{-1}(h_2\nabla)\hat{\hat{W}}^{(2)} + \left(\rho_2 - \rho_1\right)g\hat{W}^{(1)} \tag{75}$$

For solving the linear problem we need to solve for $\hat{W}^{(1)}$ using (21), then solve for $\hat{\hat{W}}^{(2)}$ using (73). After this we can use equations (72) and (75) in order to find the potentials $\hat{\hat{\Phi}}^{(2)}$ and $\hat{\Phi}^{(1)}$.

3.3 Constructing the equations for solving the Dirichlet to Neumann relation for the nonlinear problem

For the first layer we can use the same equations (15) and (45) that have been developed in section 2.4 and in the same way find $\tilde{W}^{(1)}$. For the second layer let us use (66) and (62) to obtain

$$\begin{aligned} -\left(\nabla\eta\cos\left(\bar{\eta}\nabla\right)\nabla+\sin\left(\bar{\eta}\nabla\right)\right)\hat{\tilde{\Phi}}^{(2)}- \\ \left(\nabla\eta\sin\left(\bar{\eta}\nabla\right)\nabla-\cos\left(\bar{\eta}\nabla\right)\right)\hat{\tilde{W}}^{(2)}=-\nabla\tilde{\Phi}^{(1)}\nabla\eta+\tilde{W}^{(1)}. \end{aligned} \tag{76}$$

Here $\bar{\eta}=\eta-h_2$. By using this equation with (66) on $z=h_2+\xi$ we can solve for $\hat{\tilde{\Phi}}^{(2)}$and $\hat{\tilde{W}}^{(2)}$.

3.4 Constructing equations for the nonlinear time marching

Again an analytical relation between $\tilde{\Phi}_t^{(1)}$ and $\tilde{\Phi}_t^{(2)}$must be developed in order to get equations for $\tilde{\Phi}$ using the dynamical interface boundary condition (61). Now by using (66) we get three equations relating $\tilde{\Phi}^{(2)}$, $\tilde{W}^{(2)}$ and $\tilde{\tilde{\Phi}}^{(2)}$ to $\hat{\tilde{\Phi}}^{(2)}$and $\hat{\tilde{W}}^{(2)}$. Eliminating $\hat{\tilde{\Phi}}^{(2)}$ and $\hat{\tilde{W}}^{(2)}$ from these three equations leaves a relation between $\tilde{W}^{(2)}$, $\tilde{\Phi}^{(2)}$ and $\tilde{\tilde{\Phi}}^{(2)}$

$$\tilde{W}^{(2)}=-\cot\left(\left(h_2+\xi-\eta\right)\nabla\right)\nabla\tilde{\Phi}^{(2)}+\sin^{-1}\left(\left(h_2+\xi-\eta\right)\nabla\right)\nabla\tilde{\tilde{\Phi}}^{(2)}. \tag{77}$$

By using (13) we can define $\tilde{\Phi}^{(1)}$ and $\tilde{W}^{(1)}$ with respect to $\hat{\Phi}^{(1)}$and $\hat{W}^{(1)}$together with (15) we can find a relation between $\tilde{W}^{(1)}$ and $\tilde{\Phi}^{(1)}$:

$$\tilde{W}^{(1)}=\frac{\cos\left(\left(h_1-\eta\right)\nabla\right)}{\cos\left(\left(h_1+\eta\right)\nabla\right)}\nabla\tilde{\Phi}^{(1)}. \tag{78}$$

Now let us substitute these two relations (77), (78) into the kinematic interface boundary condition (62) and derive it with respect to time to get

$$\begin{aligned} \nabla\eta_t-\eta_t\cos\left(2h_1\nabla\right)\sin^{-2}\left(\left(h_1+\eta\right)\nabla\right)\nabla\ \ \tilde{\Phi}^{(1)}- \\ \nabla\eta_t-\left(\xi_t-\eta_t\right)\sin^{-2}\left(\left(h_2+\xi-\eta\right)\nabla\right)\nabla\ \ \tilde{\Phi}^{(2)}+ \\ \left(\xi_t-\eta_t\right)\cot\left(\left(h_2+\xi-\eta\right)\nabla\right)\sin^{-1}\left(\left(h_2+\xi-\eta\right)\nabla\right)\nabla\tilde{\tilde{\Phi}}^{(2)}+ \\ \sin^{-1}\left(\left(h_2+\xi-\eta\right)\nabla\right)\tilde{\tilde{\Phi}}_t^{(2)}+ \\ \nabla\eta+\cos^{-1}\left(\left(h_1+\eta\right)\nabla\right)\sin\left(\left(h_1-\eta\right)\nabla\right)\ \ \tilde{\Phi}_t^{(1)}- \\ \left(\nabla\eta+\cot\left(\left(h_2+\xi-\eta\right)\nabla\right)\right)\tilde{\Phi}_t^{(2)}=0 \end{aligned} \tag{79}$$

Equation (79) together with equations (60) and (61) is a set of equations for $\tilde{\Phi}^{(1)}$, $\tilde{\Phi}^{(2)}$ and $\tilde{\tilde{\Phi}}^{(2)}$ and equations (2) and (59) are equations for η and ξ.

4 Internal interfacial waves between two exponentially stratified layers with a rigid-lid

4.1 An infinite series solution for the Laplace equation in each layer over a horizontal bottom

The equations governing the irrotational flow of an incompressible inviscid fluid in the lower layer over a horizontal bottom are:

$$\nabla^2\Phi^{(1)} + \Phi^{(1)}_{zz} + \frac{\frac{d\rho_1}{dz}}{\rho_1}\Phi^{(1)}_z = 0 \qquad -h_1 < z < \eta \tag{80}$$

$$\eta_t + \nabla\Phi^{(1)}\nabla\eta - \Phi^{(1)}_z = 0 \qquad z = \eta \tag{81}$$

$$\Phi^{(1)}_t + \frac{1}{2}\left|\nabla\Phi^{(1)}\right|^2 + \frac{1}{2}\left(\Phi^{(1)}_z\right)^2 + g\eta - \frac{P}{\rho_1} = 0 \qquad z = \eta \tag{82}$$

$$\Phi^{(1)}_z = 0 \qquad z = -h_1 \tag{83}$$

The equations governing the irrotational flow of an incompressible inviscid fluid in the upper layer under a rigid-lid are:

$$\nabla^2\Phi^{(2)} + \Phi^{(2)}_{zz} + \frac{\frac{d\rho_2}{dz}}{\rho_2}\Phi^{(2)}_z = 0 \qquad \eta < z < h_2 \tag{84}$$

$$\eta_t + \nabla\Phi^{(2)}\nabla\eta - \Phi^{(2)}_z = 0 \qquad z = \eta \tag{85}$$

$$\Phi^{(2)}_t + \frac{1}{2}\left|\nabla\Phi^{(2)}\right|^2 + \frac{1}{2}\left(\Phi^{(2)}_z\right)^2 + g\eta - \frac{P}{\rho_2} = 0 \qquad z = \eta \tag{86}$$

$$\Phi^{(2)}_z = 0 \qquad z = h_2 \tag{87}$$

By the use of (81), (82), (85), and (86) we can construct two interface conditions on $z = \eta$:

$$\rho_1\left(\Phi^{(1)}_t + \tfrac{1}{2}\left|\nabla\Phi^{(1)}\right|^2 + \tfrac{1}{2}\left(\Phi^{(1)}_z\right)^2 + g\eta\right) =$$

$$\rho_2\left(\Phi^{(2)}_t + \tfrac{1}{2}\left|\nabla\Phi^{(2)}\right|^2 + \tfrac{1}{2}\left(\Phi^{(2)}_z\right)^2 + g\eta\right) \tag{88}$$

$$-\nabla\Phi^{(1)}\nabla\eta + \Phi^{(1)}_z = -\nabla\Phi^{(2)}\nabla\eta + \Phi^{(2)}_z \tag{89}$$

Here Φ is the velocity potential, h_1 and h_2 are the water depths of the two layers, P is the pressure and η the interface elevation. The horizontal gradient operator relates Φ to the horizontal velocity, $\mathbf{u}$:

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right), \quad \mathbf{u} = (u, v). \tag{90}$$

For convenience we denote $\Phi_z = W$ and use the hat ($\hat{}$) and tilde ($\tilde{}$) signs to denote the value on $z = 0$ and on $z = \eta$ respectively.

For an exponentially stratified fluid the ratios $\frac{\frac{d\rho_1}{dz}}{\rho_1}$ and $\frac{\frac{d\rho_2}{dz}}{\rho_2}$ are constant and will be regarded as small parameters

$$\varepsilon_1 = \frac{\frac{d\rho_1}{dz}}{\rho_1}, \qquad \varepsilon_2 = \frac{\frac{d\rho_2}{dz}}{\rho_2}.$$

Notice that for linearly stratified fluid ε_1 and ε_2 will no longer be constants but functions of z. Nevertheless, their z derivatives will be of $O\ \ \varepsilon^2\ \ $ and every additional z derivative will further more increase the order, so if we choose to have an accuracy of $O(\varepsilon)$ the following derivations hold for linearly stratified fluids as well.

Now let us introduce an expansion of the velocity potential as a power series in the vertical coordinates with a small perturbation related to the stratification:

$$\begin{aligned}\Phi^{(1)}(x,y,z,t) &= \Phi_0^{(1)}(x,y,z,t) + \varepsilon_1\Phi_1^{(1)}(x,y,z,t) = \\ &= \textstyle\sum_{n=0}^{\infty} z^n\left(\phi_{0,n}(x,y,t) + \varepsilon_1\phi_{1,n}(x,y,t)\right).\end{aligned} \tag{91}$$

By substituting (91) into (80) and collecting equal powers of z we obtain recursive relations

$$\phi_{0,n+2}^{(1)} = -\frac{\nabla^2\phi_{0,n}^{(1)}}{(n+1)(n+2)}, \qquad \phi_{1,n+2}^{(1)} = -\frac{\nabla^2\phi_{1,n}^{(1)}}{(n+1)(n+2)} - \frac{\phi_{0,n+1}^{(1)}}{(n+2)}. \tag{92}$$

Now by substituting (92) into (91) we get

$$\begin{aligned}\Phi^{(1)}(x,y,z,t) = \sum_{n=0}^{\infty}(-1)^n \quad &\frac{z^{2n}}{(2n)!}\nabla^{2n}\phi_{0,0}^{(1)} + \frac{z^{2n+1}}{(2n+1)!}\nabla^{2n}\phi_{0,1}^{(1)} \quad + \\ \sum_{n=0}^{\infty}(-1)^n\varepsilon_1 \quad &\frac{z^{2n}}{(2n)!}\nabla^{2n}\phi_{1,0}^{(1)} - \frac{z^{2n+1}n}{(2n+1)!}\nabla^{2n}\phi_{0,0}^{(1)} + \frac{z^{2n+1}}{(2n+1)!}\nabla^{2n}\phi_{1,1}^{(1)} + \frac{z^{2n}n}{(2n)!}\nabla^{2n-2}\phi_{0,1}^{(1)}\end{aligned} \tag{93}$$

This is a series solution with 4 unknown functions $\phi_{0,0}^{(1)}$ and $\phi_{0,1}^{(1)}$, $\phi_{1,0}^{(1)}$ and $\phi_{1,1}^{(1)}$. Note that the velocities in layer 1 at the undisturbed interface $z = 0$ are given by

$$\begin{aligned}\hat{\mathbf{u}}^{(1)} &= \nabla\hat{\Phi}_0^{(1)} + \varepsilon_1\nabla\hat{\Phi}_1^{(1)} = \nabla\phi_{0,0}^{(1)} + \varepsilon_1\nabla\phi_{1,0}^{(1)}, \\ \hat{W}^{(1)} &= \hat{W}_0^{(1)} + \varepsilon_1\hat{W}_1^{(1)} = \phi_{0,1}^{(1)} + \varepsilon_1\phi_{1,1}^{(1)}. \qquad @z = 0\end{aligned} \tag{94}$$

By using Rayleigh's notation and (94) we can write equation (93) as

$$\begin{aligned}\Phi^{(1)}(x,y,z,t) = \quad &\cos(z\nabla) + \tfrac{1}{2}\varepsilon_1 z\cos(z\nabla) + \varepsilon_1\tfrac{\sin(z\nabla)}{2\nabla} \quad \hat{\Phi}_0^{(1)} + \\ &\tfrac{\sin(z\nabla)}{\nabla} - \varepsilon_1\tfrac{z\sin(z\nabla)}{2\nabla} \quad \hat{W}_0^{(1)} + \varepsilon_1\cos(z\nabla)\hat{\Phi}_1^{(1)} + \varepsilon_1\tfrac{\sin(z\nabla)}{\nabla}\hat{W}_1^{(1)}\end{aligned} \tag{95}$$

Equation (95) enables us to write $\nabla\Phi^{(1)}$ and $W^{(1)}$:

$$\nabla\Phi^{(1)} = \quad 1+\tfrac{1}{2}\varepsilon_1 z \;\; \cos(z\nabla)\nabla + \tfrac{1}{2}\varepsilon_1 \sin(z\nabla) \;\; \hat{\Phi}_0^{(1)} +$$
$$1-\tfrac{1}{2}\varepsilon_1 z \;\; \sin(z\nabla)\hat{W}_0^{(1)} + \varepsilon_1 \cos(z\nabla)\nabla\hat{\Phi}_1^{(1)} + \varepsilon_1 \sin(z\nabla)\hat{W}_1^{(1)} \tag{96}$$

$$W^{(1)} = \;\; \varepsilon_1 \cos(z\nabla) - \;\; 1+\tfrac{1}{2}\varepsilon_1 z \;\; \sin(z\nabla)\nabla \;\; \hat{\Phi}_0^{(1)} +$$
$$1-\tfrac{1}{2}\varepsilon_1 z \;\; \cos(z\nabla) - \tfrac{1}{2}\varepsilon_1 \tfrac{\sin(z\nabla)}{\nabla} \;\; \hat{W}_0^{(1)} - \varepsilon_1 \sin(z\nabla)\nabla\hat{\Phi}_1^{(1)} + \varepsilon_1 \cos(z\nabla)\hat{W}_1^{(1)} \tag{97}$$

Now with the use of (97) and (94) the horizontal bottom condition (83) can be expressed as

$$\begin{aligned} O \;\; \varepsilon_1^0 \;\; : \quad & \hat{W}_0^{(1)} = -\tan(h_1\nabla)\nabla\hat{\Phi}_0^{(1)} \\ O \;\; \varepsilon_1^1 \;\; : \quad & \hat{W}_1^{(1)} = -\tan(h_1\nabla)\nabla\hat{\Phi}_1^{(1)} - \tfrac{1}{2}h_1\tan^2(h_1\nabla)\hat{\Phi}_0^{(1)} \end{aligned} \tag{98}$$

A similar relation can be found for the upper layer

$$\begin{aligned} O \;\; \varepsilon_2^0 \;\; : \quad & \hat{W}_0^{(2)} = \tan(h_2\nabla)\nabla\hat{\Phi}_0^{(2)} \\ O \;\; \varepsilon_2^1 \;\; : \quad & \hat{W}_1^{(2)} = \tan(h_2\nabla)\nabla\hat{\Phi}_1^{(2)} + \tfrac{1}{2}h_2\tan^2(h_2\nabla)\hat{\Phi}_0^{(2)} \end{aligned} \tag{99}$$

4.2 Constructing the accurate equations for the linear problem

For small amplitude waves in the sense that the ration between the amplitude to the wave length (wave slope) is small the problem becomes linear. We can see that the kinematic and dynamic interface conditions (89) and (88) become

$$\hat{W}^{(1)} = \hat{W}^{(2)} \qquad @z=0 \tag{100}$$

$$\rho_1 \;\; \hat{\Phi}_t^{(1)} + g\eta \;\; = \rho_2 \;\; \hat{\Phi}_t^{(2)} + g\eta \qquad @z=0 \tag{101}$$

By the use of (98), (99) and (100) we can construct relations between $\hat{\Phi}_0^{(1)}$, $\hat{\Phi}_1^{(1)}$, $\hat{\Phi}_0^{(2)}$ and $\hat{\Phi}_1^{(2)}$

$$\hat{\Phi}_0^{(2)} = -\tfrac{\tan(h_1\nabla)}{\tan(h_2\nabla)}\hat{\Phi}_0^{(1)},$$
$$\hat{\Phi}_1^{(2)} = \tfrac{h_2\tan(h_1\nabla)}{2\nabla}\hat{\Phi}_0^{(1)} - \tfrac{\varepsilon_1}{\varepsilon_2} \;\; \tfrac{\tan(h_1\nabla)}{\tan(h_2\nabla)}\hat{\Phi}_1^{(1)} + \tfrac{h_1\tan^2(h_1\nabla)}{2\tan(h_2\nabla)\nabla}\hat{\Phi}_0^{(1)} \qquad @z=0 \tag{102}$$

By using (102), the linear version of (82) and the time derivative of (101) we get

$$O \;\; \varepsilon_1^0 \;\; : \qquad \rho_1 + \rho_2\tfrac{\tan(h_1\nabla)}{\tan(h_2\nabla)} \;\; \tfrac{\partial^2}{\partial t^2}\hat{\Phi}_0^{(1)} = (\rho_2-\rho_1)\, g\hat{W}_0^{(1)}$$

$$\begin{aligned} O \;\; \varepsilon_1^1 \;\; : \quad & \rho_1 + \rho_2\tfrac{\varepsilon_1}{\varepsilon_2}\tfrac{\tan(h_1\nabla)}{\tan(h_2\nabla)} \;\; \tfrac{\partial^2}{\partial t^2}\hat{\Phi}_1^{(1)} = \\ & (\rho_2-\rho_1)\, g\hat{W}_1^{(1)} - \;\; \tfrac{h_2\tan(h_1\nabla)}{2\nabla} + \tfrac{\varepsilon_1}{\varepsilon_2}\tfrac{h_1\tan^2(h_1\nabla)}{2\tan(h_2\nabla)\nabla} \;\; \tfrac{\partial^2}{\partial t^2}\hat{\Phi}_0^{(1)} \end{aligned} \tag{103}$$

Now by using equations (103) and (98) we can solve $\hat{\Phi}_0^{(1)}$, $\hat{\Phi}_1^{(1)}$, $\hat{W}_0^{(1)}$ and $\hat{W}_1^{(1)}$.

4.3 Constructing approximate equations for the nonlinear Dirichlet to Neumann problem

Equation (80) allows us to replace horizontal differentiation by vertical ones in order to obtain

$$\nabla^{2m}\Phi^{(1)} = (-1)^m \quad \frac{\partial^{2m}\Phi^{(1)}}{\partial z^{2m}} + m\varepsilon\frac{\partial^{2m-1}\Phi^{(1)}}{\partial z^{2m-1}}\Bigg) \tag{104}$$

By extending one of the main Boussinesq concepts we look for $\Phi^{(1)}$ of the form

$$\Phi^{(1)}(x,y,z,t) = \sum_{n=0}^{\infty}(z+h_1)^n\left(\phi_{0,n}(x,y,t) + \varepsilon_1\phi_{1,n}(x,y,t)\right). \tag{105}$$

From (105) it follows that

$$\frac{\partial^m\Phi^{(1)}}{\partial z^m} = \sum_{n=m}^{\infty}\frac{n!}{(n-m)!}(z+h_1)^{n-m}\left(\phi_{0,n}(x,y,t) + \varepsilon_1\phi_{1,n}(x,y,t)\right). \tag{106}$$

Now by using (104) and (106) we get

$$\nabla^{2m}\Phi^{(1)} = (-1)^m\textstyle\sum_{n=2m}^{\infty}\frac{n!}{(n-2m)!}(z+h_1)^{n-2m}\phi_{0,n} + \tag{107}$$
$$\varepsilon_1(-1)^m \quad \textstyle\sum_{n=2m}^{\infty}\frac{n!}{(n-2m)!}(z+h_1)^{n-2m}\phi_{1,n} + m\sum_{n=2m-1}^{\infty}\frac{n!}{(n-2m+1)!}(z+h_1)^{n-2m+1}\phi_{0,n} \quad + O\ \ \varepsilon_1^2\ \ ,$$

$$\nabla^{2m}W^{(1)} = (-1)^m\textstyle\sum_{n=2m+1}^{\infty}\frac{n!}{(n-2m-1)!}(z+h_1)^{n-2m-1}\phi_{0,n} + \tag{108}$$
$$\varepsilon_1(-1)^m \quad \textstyle\sum_{n=2m+1}^{\infty}\frac{n!}{(n-2m-1)!}(z+h_1)^{n-2m-1}\phi_{1,n} + m\sum_{n=2m}^{\infty}\frac{n!}{(n-2m)!}(z+h_1)^{n-2m}\phi_{0,n} \quad + O\ \ \varepsilon_1^2\ \ .$$

Next by substituting $m = 0, 1$ into equations (107) and (108) we get the following equations for $\hat{\Phi}^{(1)}$, $\nabla^2\hat{\Phi}^{(1)}$, $\hat{W}^{(1)}$ and $\nabla^2\hat{W}^{(1)}$:

$$O\ \ \varepsilon_1^0\ \ : \qquad \begin{aligned} \hat{\Phi}_0^{(1)} &= \textstyle\sum_{n=0}^{3} h_1^n\phi_{0,n}, \\ \hat{W}_0^{(1)} &= \textstyle\sum_{n=1}^{3} nh_1^{n-1}\phi_{0,n}, \\ \nabla^2\hat{W}_0^{(1)} &= -3!\phi_{0,3}, \\ \nabla^2\hat{\Phi}_0^{(1)} &= -\textstyle\sum_{n=2}^{3}\frac{n!}{(n-2)!}h_1^{n-2}\phi_{0,n}. \qquad @z=0 \end{aligned} \tag{109}$$

$$O\ \ \varepsilon_1^1\ \ : \qquad \begin{aligned} \hat{\Phi}_1^{(1)} &= \textstyle\sum_{n=0}^{3} h_1^n\phi_{1,n}, \\ \hat{W}_1^{(1)} &= \textstyle\sum_{n=1}^{3} nh_1^{n-1}\phi_{1,n}, \\ \nabla^2\hat{\Phi}_1^{(1)} &= -\textstyle\sum_{n=2}^{3}\frac{n!}{(n-2)!}h_1^{n-2}\phi_{1,n} - \sum_{n=1}^{3}\frac{n!}{(n-1)!}h_1^{n-1}\phi_{0,n}, \\ \nabla^2\hat{W}_1^{(1)} &= -3!\phi_{1,3} - \textstyle\sum_{n=2}^{3}\frac{n!}{(n-2)!}h_1^{n-2}\phi_{0,n}. \qquad @z=0 \end{aligned} \tag{110}$$

Notice that (107) and (108) were truncated. Only the first 4 base functions were used in each order. This enables us to find these base functions using the 8 equations written in (109) and (110). Had we wanted to use more base functions for higher accuracy, we would have needed to increase the number of equations by substituting $m = 2$, $m = 3$ and so on into equations (107) and (108). That would also have increased the order of the derivatives, which would have been more complex to solve.

Note that the accuracy of this calculation is lower than the one presented in section 2.4 because we need to use the odd base functions as well as the even ones. Thus, for the same number of base functions the expansion of $\Phi^{(1)}$ shown in equation (105) gives us a power expansion up to z^3, whereas in section 2.4 it was up to z^6.

By solving equation sets (109) and (110) for the base functions $\phi^{(1)}_{0,n}$, $n = 0..3$ and $\phi^{(1)}_{1,n}$, $n = 0..3$ in terms of $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$ and $\hat{W}^{(1)}_1$ and their second horizontal derivatives we get

$$
\begin{aligned}
O\ \ \varepsilon_1^0\ :\qquad \phi^{(1)}_{0,0} &= \hat{\Phi}^{(1)}_0 - h_1\hat{W}^{(1)}_0 - \tfrac{1}{2}\nabla^2\hat{\Phi}^{(1)}_0 + \tfrac{1}{6}\nabla^2\hat{W}^{(1)}_0 \\
\phi^{(1)}_{0,1} &= \hat{W}^{(1)}_0 + h_1\nabla^2\hat{\Phi}^{(1)}_0 - \tfrac{1}{2}h_1^2\nabla^2\hat{W}^{(1)}_0 \\
\phi^{(1)}_{0,2} &= -\tfrac{1}{2}\nabla^2\hat{\Phi}^{(1)}_0 + \tfrac{1}{2}h_1\nabla^2\hat{W}^{(1)}_0 \\
\phi^{(1)}_{0,3} &= -\tfrac{1}{6}\nabla^2\hat{W}^{(1)}_0
\end{aligned}
\tag{111}
$$

$$
\begin{aligned}
O\ \ \varepsilon_1^1\ :\quad \phi^{(1)}_{1,0} &= \hat{\Phi}^{(1)}_1 - h_1\hat{W}^{(1)}_1 - \tfrac{1}{6}h_1^2\nabla^2\hat{\Phi}^{(1)}_0 + \tfrac{1}{6}h_1^2\nabla^2\hat{W}^{(1)}_0 - \tfrac{1}{2}h_1^2\nabla^2\hat{\Phi}^{(1)}_1 - \tfrac{1}{2}h_1^2\hat{W}^{(1)}_0 \\
\phi^{(1)}_{1,1} &= \hat{W}^{(1)}_1 + h_1\nabla^2\hat{\Phi}^{(1)}_1 + h_1\hat{W}^{(1)}_0 - \tfrac{1}{2}h_1\nabla^2\hat{W}^{(1)}_0 + \tfrac{1}{2}h_1\nabla^2\hat{\Phi}^{(1)}_0 \\
\phi^{(1)}_{1,2} &= -\tfrac{1}{2}\nabla^2\hat{\Phi}^{(1)}_0 + \tfrac{1}{2}\nabla^2\hat{W}^{(1)}_0 - \tfrac{1}{2}\hat{W}^{(1)}_0 - \tfrac{1}{2}\nabla^2\hat{\Phi}^{(1)}_1 \\
\phi^{(1)}_{1,3} &= -\tfrac{1}{6h_1}\nabla^2\hat{W}^{(1)}_0 + \tfrac{1}{6h_1}\nabla^2\hat{\Phi}^{(1)}_0
\end{aligned}
\tag{112}
$$

Substituting the solutions (111) and (112) into (91) on $z = \eta$ gives us relations involving $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$, $\hat{W}^{(1)}_1$, $\tilde{\Phi}^{(1)}_0$ and $\tilde{\Phi}^{(1)}_1$:

$$
\tilde{\Phi}^{(1)}_0 = \hat{\Phi}^{(1)}_0 - (h_1-\eta)\,\hat{W}^{(1)}_0 - \tfrac{1}{2}(h_1-\eta)^2\,\nabla^2\hat{\Phi}^{(1)}_0 + \tfrac{1}{6}(h_1-\eta)^3\,\nabla^2\hat{W}^{(1)}_0 \tag{113}
$$

$$
\begin{aligned}
\tilde{\Phi}^{(1)}_1 = \hat{\Phi}^{(1)}_1 - (h_1-\eta)\,\hat{W}^{(1)}_1 - \tfrac{1}{2}(h_1-\eta)^2\,\nabla^2\hat{\Phi}^{(1)}_1 - \\
-\tfrac{1}{2}(h_1-\eta)^2\,\hat{W}^{(1)}_0 - \tfrac{1}{6h_1}(h_1-\eta)^3\,\nabla^2\hat{\Phi}^{(1)}_0 + \tfrac{1}{6h_1}(h_1-\eta)^3\,\nabla^2\hat{W}^{(1)}_0
\end{aligned}
\tag{114}
$$

Labels (98), (113) and (114) presents us with 4 equations for $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$ and $\hat{W}^{(1)}_1$. We can solve $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$ and $\hat{W}^{(1)}_1$ using a numerical method such as finite differences. By implementing this method we receive 2 sets of linear equations, which take the form of

$$\begin{bmatrix} A^{(1)}_{0\,\text{padé}} & -B^{(1)}_{0\,\text{padé}} \\ A^{(1)}_{0\,\text{nl}} & B^{(1)}_{0\,\text{nl}} \end{bmatrix} \begin{bmatrix} \hat{\Phi}^{(1)}_0 \\ \hat{W}^{(1)}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \tilde{\Phi}^{(1)}_0 \end{bmatrix}, \tag{115}$$

$$\begin{bmatrix} A^{(1)}_{1\,\text{padé}} & -B^{(1)}_{1\,\text{padé}} \\ A^{(1)}_{1\,\text{nl}} & B^{(1)}_{1\,\text{nl}} \end{bmatrix} \begin{bmatrix} \hat{\Phi}^{(1)}_1 \\ \hat{W}^{(1)}_1 \end{bmatrix} = \begin{bmatrix} C^{(1)}_1 \hat{\Phi}^{(1)}_0 \\ \tilde{\Phi}^{(1)}_1 + C^{(1)}_2 \hat{\Phi}^{(1)}_0 + C^{(1)}_3 \hat{W}^{(1)}_0 \end{bmatrix}. \tag{116}$$

Here $A^{(1)}_{\text{padé}}$ and $B^{(1)}_{\text{padé}}$ are the finite difference matrices representing the Dirichlet to Neumann relation (label 98), $A^{(1)}_{\text{nl}}$and $B^{(1)}_{\text{nl}}$are the finite difference matrices representing the relation between the undisturbed potential and vertical velocity and the potential on the interfacial wave. $C^{(1)}_1$ and $C^{(1)}_2$ are the finite difference matrices representing the relation between $\hat{W}^{(1)}_1$ and $\hat{\Phi}^{(1)}_1$ to $\hat{\Phi}^{(1)}_0$. $C^{(1)}_3$ is the finite difference matrix representing the relation between $\hat{W}^{(1)}_1$ and $\hat{\Phi}^{(1)}_1$ to $\hat{W}^{(1)}_0$. By using $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$ and $\hat{W}^{(1)}_1$ we can find the base functions using (111) and (112), which gives us the properties of the flow at every point in the layer.

Now let us present a way to calculate the properties of the flow in the upper layer. We can easily find $\tilde{W}^{(2)}_0$ and $\tilde{W}^{(2)}_1$ using (89). Applying the same technique described in subsection 2.1 we can construct an equivalent version of equations (95) and (97) for the upper layer

$$\begin{aligned} \Phi^{(2)} = \quad & \cos(z\nabla) + \tfrac{1}{2}\varepsilon_2 z\cos(z\nabla) + \varepsilon_2 \tfrac{\sin(z\nabla)}{2\nabla} \quad \Phi^{(2)}_0 + \\ & \tfrac{\sin(z\nabla)}{\nabla} - \varepsilon_2 \tfrac{z\sin(z\nabla)}{2\nabla} \quad W^{(2)}_0 + \varepsilon_2\cos(z\nabla)\Phi^{(2)}_1 + \varepsilon_2 \tfrac{\sin(z\nabla)}{\nabla} W^{(2)}_1, \end{aligned} \tag{117}$$

$$\begin{aligned} W^{(2)} = \quad & -1 + \tfrac{1}{2}\varepsilon_2 z \quad \sin(z\nabla)\nabla\Phi^{(2)}_0 + \quad 1 - \tfrac{1}{2}\varepsilon_2 z \quad \cos(z\nabla) - \tfrac{1}{2}\varepsilon_2 z \tfrac{\sin(z\nabla)}{\nabla} \quad W^{(2)}_0 - \\ & \varepsilon_2 \sin(z\nabla)\nabla\Phi^{(2)}_1 + \varepsilon_2\cos(z\nabla) W^{(2)}_1 \end{aligned} \tag{118}$$

On $z = \eta$ we can use equations (117) and (118) to solve $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$ and $\hat{W}^{(1)}_1$ and substitute the solution into (117) and (118) to get $\Phi^{(2)}$ and $W^{(2)}$ at every point in the layer.

4.4 Constructing the equations for the nonlinear time marching

An analytical relation between $\tilde{\Phi}^{(1)}_t$ and $\tilde{\Phi}^{(2)}_t$ must be developed in order to enable marching $\tilde{\Phi}$ in time using the dynamical interface boundary condition (88). Equations (95), (97) and (98) on $z = \eta$ are a set of 6 equations relating $\hat{\Phi}^{(1)}_0$, $\hat{W}^{(1)}_0$, $\hat{\Phi}^{(1)}_1$, $\hat{W}^{(1)}_1$, $\tilde{\Phi}^{(1)}_0$, $\tilde{W}^{(1)}_0$, $\tilde{\Phi}^{(1)}_1$ and $\tilde{W}^{(1)}_1$. By eliminating these 6 equations we get

$$\tilde{W}^{(1)}_0 = -\tan\left((h_1+\eta)\nabla\right)\nabla\tilde{\Phi}^{(1)}_0 \tag{119}$$

$$\begin{aligned} \tilde{W}^{(1)}_1 = & -\tan\left((h_1+\eta)\nabla\right)\nabla\tilde{\Phi}^{(1)}_1 + \\ & \tfrac{1}{4}\sec^2\left((h_1+\eta)\nabla\right)\left(1 - h_1 + (2+h_1)\cos\left(2h_1\nabla\right) + 2\cos\left(2\eta\nabla\right) + 2\eta\sin\left(2h_1\nabla\right)\nabla\right)\tilde{\Phi}^{(1)}_0 \end{aligned}$$

Similar relations can be found for the upper layer

$$\tilde{W}_0^{(2)} = \tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\tilde{\Phi}_0^{(2)} \tag{120}$$
$$\tilde{W}_1^{(2)} = \tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\tilde{\Phi}_1^{(2)} +$$
$$\tfrac{1}{4}\sec^2\left(\left(h_2-\eta\right)\nabla\right)\left(1-h_2+\left(2+h_2\right)\cos\left(2h_2\nabla\right)+2\cos\left(2\eta\nabla\right)-2\eta\sin\left(2h_2\nabla\right)\nabla\right)\tilde{\Phi}_0^{(2)}$$

By using equations (119), (120) and (89) we get relations between $\tilde{\Phi}_0^{(1)}$, $\tilde{\Phi}_1^{(1)}$, $\tilde{\Phi}_0^{(2)}$and $\tilde{\Phi}_1^{(2)}$

$$-\nabla\eta\tilde{\Phi}_0^{(1)} - \tan\left(\left(h_1+\eta\right)\nabla\right)\nabla\tilde{\Phi}_0^{(1)} = \nabla\eta\tilde{\Phi}_0^{(2)} + \tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\tilde{\Phi}_0^{(2)}$$

$$\tfrac{1}{4}\sec^2\left(\left(h_1+\eta\right)\nabla\right)\left(1-h_1+\left(2+h_1\right)\cos\left(2h_1\nabla\right)+2\cos\left(2\eta\nabla\right)+2\eta\sin\left(2h_1\nabla\right)\nabla\right)\tilde{\Phi}_0^{(1)} -$$
$$\nabla\eta\tilde{\Phi}_1^{(1)} - \tan\left(\left(h_1+\eta\right)\nabla\right)\nabla\tilde{\Phi}_1^{(1)} =$$
$$\tfrac{1}{4}\sec^2\left(\left(h_2-\eta\right)\nabla\right)\left(1-h_2+\left(2+h_2\right)\cos\left(2h_2\nabla\right)+2\cos\left(2\eta\nabla\right)-2\eta\sin\left(2h_2\nabla\right)\nabla\right)\tilde{\Phi}_0^{(2)} -$$
$$\nabla\eta\tilde{\Phi}_1^{(2)} + \tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\tilde{\Phi}_1^{(2)} \tag{121}$$

By taking the time derivative of (121) we get

$$\eta_t\sec^2\left(\left(h_1+\eta\right)\nabla\right)\nabla^2\tilde{\Phi}_0^{(1)} - \left(\nabla\eta+\tan\left(\left(h_1+\eta\right)\nabla\right)\nabla\right)\tfrac{\partial}{\partial t}\tilde{\Phi}_0^{(1)} =$$
$$\eta_t\sec^2\left(\left(h_2-\eta\right)\nabla\right)\nabla^2\tilde{\Phi}_0^{(2)} - \left(\nabla\eta-\tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\right)\tfrac{\partial}{\partial t}\tilde{\Phi}_0^{(2)}$$

$$-\eta_t\sec^2\left(\left(h_1+\eta\right)\nabla\right)\nabla^2\tilde{\Phi}_1^{(1)} + \tfrac{1}{2}\eta_t\sec^2\left(\left(h_1+\eta\right)\nabla\right)$$
$$\left(-\sin\left(2\eta\nabla\right)+\left(1-h_1+\left(2+h_1\right)\cos\left(2h_1\nabla\right)\right)\tan\left(\left(h_1+\eta\right)\nabla\right)+2\sin\left(2h_1\nabla\right)\right)\nabla\tilde{\Phi}_0^{(1)} +$$
$$\left(\cos\left(2\eta\nabla\right)+2\nabla\eta\sin\left(2h_1\nabla\right)\right)\tan\left(\left(h_1+\eta\right)\nabla\right)\nabla\tilde{\Phi}_0^{(1)} +$$
$$\tfrac{1}{4}\sec^2\left(\left(h_1+\eta\right)\nabla\right)\left(1-h_1+\left(2+h_1\right)\cos\left(2h_1\nabla\right)+2\cos\left(2\eta\nabla\right)+2\eta\sin\left(2h_1\nabla\right)\nabla\right)\tfrac{\partial}{\partial t}\tilde{\Phi}_0^{(1)} -$$
$$\left(\nabla\eta+\tan\left(\left(h_1+\eta\right)\nabla\right)\nabla\right)\tfrac{\partial}{\partial t}\tilde{\Phi}_1^{(1)} =$$
$$-\eta_t\sec^2\left(\left(h_2-\eta\right)\nabla\right)\nabla^2\tilde{\Phi}_1^{(2)} + \tfrac{1}{2}\eta_t\sec^2\left(\left(h_2-\eta\right)\nabla\right)$$
$$\left(-\sin\left(2\eta\nabla\right)+\left(1-h_2+\left(2+h_2\right)\cos\left(2h_2\nabla\right)\right)\tan\left(\left(h_2-\eta\right)\nabla\right)-2\sin\left(2h_2\nabla\right)\right)\nabla\tilde{\Phi}_0^{(2)}$$
$$\left(\cos\left(2\eta\nabla\right)+2\nabla\eta\sin\left(2h_2\nabla\right)\right)\tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\tilde{\Phi}_0^{(2)} +$$
$$\tfrac{1}{4}\sec^2\left(\left(h_2-\eta\right)\nabla\right)\left(1-h_2+\left(2+h_2\right)\cos\left(2h_2\nabla\right)+2\cos\left(2\eta\nabla\right)-2\eta\sin\left(2h_2\nabla\right)\nabla\right)\tfrac{\partial}{\partial t}\tilde{\Phi}_0^{(2)} -$$
$$\left(\nabla\eta-\tan\left(\left(h_2-\eta\right)\nabla\right)\nabla\right)\tfrac{\partial}{\partial t}\tilde{\Phi}_1^{(2)} \tag{122}$$

Now by using equations (122), (81) and (88) we can march the problem in time.

5 Surface Quasi-Geostrophic Model

The quasi-geostrophic equation for constant stratification is

$$\frac{\partial}{\partial t} + \Psi_x \frac{\partial}{\partial y} - \Psi_y \frac{\partial}{\partial x} \qquad \nabla^2 \Psi + \frac{1}{S} \Psi_{zz} + \beta y \quad = 0 \tag{123}$$

$$\mathbf{u} = (-\Psi_y, \Psi_x), \qquad \Psi_z = \Theta \tag{124}$$

Here $S = \frac{N^2 H^2}{f_0^2 L^2}$, N is the buoyancy frequency, f_0 is the rotation frequency, H is the vertical length scale, L is the horizontal length scale and Ψ is the horizontal stream function and its z-derivative relates to the potential temperature Θ. Assuming zero initial potential vorticity equation (123) becomes

$$\nabla^2 \Psi + \frac{1}{S} \Psi_{zz} + \beta y = 0 \tag{125}$$

The Boundary conditions for this one layer model are

$$\Theta_t + \Psi_x \Theta_y - \Psi_y \Theta_x = 0 \qquad @z = 0 \tag{126}$$

$$\Psi_z = 0 \qquad @z = -H \tag{127}$$

Equations (125), (126) and (127) represents the surface quasi-geostrophic model. We introduce an expansion of Ψ as a power series in z:

$$\Psi(x, y, z, t) = \sum_{n=0}^{\infty} z^n \psi_n(x, y, t) - \frac{1}{6} \beta y^3 \tag{128}$$

By substituting equation (128) into equation (125) and using equation (127) we get

$$\cos(\sqrt{S} H \nabla) \hat{\Theta} + \sin(\sqrt{S} H \nabla) \nabla \hat{\Psi} = 0 \tag{129}$$

Therefore,

$$\begin{aligned} \hat{\Psi}_x &= -\cot \quad \sqrt{S} H \tfrac{\partial}{\partial x} \quad \hat{\Theta} \\ \hat{\Psi}_y &= -\cot \quad \sqrt{S} H \tfrac{\partial}{\partial y} \quad \hat{\Theta} \end{aligned} \tag{130}$$

As before, the hat sign ($\hat{}$) denotes the value on $z = 0$. Now by applying (130) to (126) we get

$$\Theta_t - \Theta_y \cot \quad \sqrt{S} H \frac{\partial}{\partial x} \quad \Theta + \Theta_x \cot \quad \sqrt{S} H \frac{\partial}{\partial y} \quad \Theta = 0 \qquad @z = 0 \tag{131}$$

The infinite differential operators in equation (131) can now be approximated using Taylor approximation or Padé approximation to any required order and solved for Θ.

6 A new Layered Quasi-Geostrophic Model

6.1 Formulating the equations for the layered model

The surface quasi-geostrophic model regards the entire water depth as one stratified layer. In this new model the water depth is divided to two stratified layers with a density jump in the interface. This represents better the ocean's density structure and should give more accurate results than the one-layered representation. Due to the density jump in the interface there is a mechanism for internal interfacial Rossby waves to propagate. The equations governing the quasi-geostrophical flow of an incompressible inviscid fluid in the lower layer over a horizontal bottom are:

$$\nabla^2\Psi^{(1)} + \frac{1}{S_1}\Psi^{(1)}_{zz} + \beta y = 0 \tag{132}$$

$$\hat{\Theta}^{(1)}_t + \hat{\mathbf{u}}^{(1)} \cdot \nabla\hat{\Theta}^{(1)} = N_1^2\hat{W}^{(1)} \qquad @z = 0 \tag{133}$$

$$\eta_t + \hat{\mathbf{u}}^{(1)} \cdot \nabla\eta = \hat{W}^{(1)} \qquad @z = 0 \tag{134}$$

$$\Theta^{(1)} = 0 \qquad @z = -h_1 \tag{135}$$

The equations governing the quasi-geostrophical flow of an incompressible inviscid fluid in the upper layer under a rigid lid are:

$$\nabla^2\Psi^{(2)} + \frac{1}{S_2}\Psi^{(2)}_{zz} + \beta y = 0 \tag{136}$$

$$\hat{\hat{\Theta}}^{(2)}_t + \hat{\mathbf{u}}^{(2)} \cdot \nabla\hat{\hat{\Theta}}^{(2)} = F \qquad @z = h_2 \tag{137}$$

$$\hat{\Theta}^{(2)}_t + \hat{\mathbf{u}}^{(2)} \cdot \nabla\hat{\Theta}^{(2)} = N_2^2\hat{W}^{(2)} \qquad @z = 0 \tag{138}$$

$$\eta_t + \hat{\mathbf{u}}^{(2)} \cdot \nabla\eta = \hat{W}^{(2)} \qquad @z = 0 \tag{139}$$

Here we added a surface potential temperature forcing function F to the surface boundary condition that is presented in equation (137). Using (133), (134), (138) and (139) we can write the potential temperature interfacial equation

$$\begin{aligned} N_2^2 \;\; \hat{\Theta}^{(1)}_t + \hat{\mathbf{u}}^{(1)} \cdot \nabla \;\; \hat{\Theta}^{(1)} - N_1^2\eta \quad &= \\ N_1^2 \;\; \hat{\Theta}^{(2)}_t + \hat{\mathbf{u}}^{(2)} \cdot \nabla \;\; \hat{\Theta}^{(2)} - N_2^2\eta \qquad & @z = 0 \end{aligned} \tag{140}$$

Here,

$$\nabla = \;\; \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \tag{141}$$

$$\mathbf{u} = (-\Psi_y, \Psi_x)\,, \qquad \Psi_z = \Theta. \tag{142}$$

Now in order for the problem to be well-posed we need to construct a pressure interfacial condition. Let us write the pressure formulation for each layer,

$$-p_s - p_2(z) = \rho_1\, g\,(h_2 - z) - \Delta\rho\, g\,(h_2 - z) - p_2^{'} \quad @\ 0 < z < h_2 \tag{143}$$

$$-p_s - p_1(z) = \rho_1\, g\,(h_2 - z) - \Delta\rho\, g\,(h_2 - \eta) - p_1^{'} \quad @\ -h_1 < z < 0. \tag{144}$$

Here $p_1^{'}$ and $p_2^{'}$ are the non-hydrostatic part of the pressure, p_s is the rigid-lid surface pressure and $\Delta\rho = \rho_2 - \rho_1$. By defining Ψ with p_s as the datum and using equations (143) and (144) we can write the pressure interfacial equation,

$$\hat{\Psi}^{(2)} + \Delta\rho\, g\eta = \hat{\Psi}^{(1)} \quad @z = 0. \tag{145}$$

By using (133) and (134) we can also construct an equation relating η to the potential temperature,

$$\eta_t = \tfrac{1}{N_1^2}\ \ \hat{\Theta}_t^{(1)} + \hat{\mathbf{u}}^{(1)} \cdot\ \ \hat{\Theta}^{(1)} - N_1^2\eta \qquad @z = 0. \tag{146}$$

6.2 An infinite series solution for the quasi-geostrophic equation in each layer over a horizontal bottom

We introduce an expansion of Ψ as a power series in z:

$$\Psi(x, y, z, t) = \sum_{n=0}^{\infty} z^n \psi_n(x, y, t) - \frac{1}{6}\beta y^3 \tag{147}$$

By substituting equation (147) for the lower layer into equation (132) and using equation (135) we get

$$\cos(\sqrt{S_1}h_1\nabla)\hat{\Theta}^{(1)} + \sin(\sqrt{S_1}h_1\nabla)\nabla\hat{\Psi}^{(1)} = 0 \tag{148}$$

and by substituting equation (147) for the upper layer into equation (136) we get

$$\Psi^{(2)} = \tfrac{\sin\left(\sqrt{S_2}(z-h_2)\nabla\right)}{\nabla}\hat{\hat{\Theta}}^{(2)} + \cos\ \ \sqrt{S_2}\,(z - h_2)\,\nabla\ \ \hat{\hat{\Psi}}^{(2)} - \tfrac{1}{6}\beta y^3. \tag{149}$$

Now by using equation (149) we can formulate the relations between the flow properties of the upper layer on the interface and the flow properties on the surface,

$$\hat{\Theta}^{(2)} = \sqrt{S_2}\cos\ \ \sqrt{S_2}h_2\nabla\ \ \hat{\hat{\Theta}}^{(2)} + \sqrt{S_2}\sin\ \ \sqrt{S_2}h_2\nabla\ \ \nabla\hat{\hat{\Psi}}^{(2)} \tag{150}$$

$$\hat{\Psi}^{(2)} = -\tfrac{\sin\left(\sqrt{S_2}h_2\nabla\right)}{\nabla}\hat{\hat{\Theta}}^{(2)} + \cos\ \ \sqrt{S_2}h_2\nabla\ \ \hat{\hat{\Psi}}^{(2)} - \tfrac{1}{6}\beta y^3. \tag{151}$$

Let us now eliminate equations (145), (148), (150) and (151) to give us an equation containing only the potential temperature functions $\hat{\Theta}^{(1)}$,$\hat{\Theta}^{(2)}$and $\hat{\hat{\Theta}}^{(2)}$,

$$\sqrt{S_2}\tfrac{1}{\nabla}\cot\ \sqrt{S_1}h_1\nabla\ \hat{\Theta}^{(1)} + \tfrac{1}{\nabla}\cot\ \sqrt{S_2}h_2\nabla\ \hat{\Theta}^{(2)} - \sqrt{S_2}\tfrac{1}{\nabla}\csc\ \sqrt{S_2}h_2\nabla\ \hat{\hat{\Theta}}^{(2)} + \sqrt{S_2}\ \Delta\rho\, g\eta - \tfrac{1}{6}\beta y^3\ = 0. \tag{152}$$

The next step is to take the derivative of equation (152) with respect to time and also use equations (137) and (146) in order to get

$$\tfrac{1}{\nabla}\cot\ \sqrt{S_2}h_2\nabla\ \hat{\Theta}_t^{(2)} - \tfrac{1}{\nabla}\csc\ \sqrt{S_2}h_2\nabla\quad F - \hat{\mathbf{u}}^{(2)}\cdot\nabla\hat{\hat{\Theta}}^{(2)}\ =$$
$$-\tfrac{\Delta\rho\, f}{N_1^2}\hat{\mathbf{u}}^{(1)}\cdot\nabla\ \hat{\Theta}^{(1)} - N_1^2\eta\ -\ \tfrac{1}{\nabla}\cot\ \sqrt{S_1}h_1\nabla\ + \tfrac{\Delta\rho\, f}{N_1^2}\ \hat{\Theta}_t^{(1)}. \tag{153}$$

At this point we have all the equations needed in order to construct the method. From initial conditions or from a prior time step we know $\hat{\Theta}^{(1)}$, $\hat{\hat{\Theta}}^{(2)}$ and η. By the use of equations (145), (148) and (151) we can find $\hat{\Psi}^{(1)}$, $\hat{\Psi}^{(2)}$ and $\hat{\hat{\Psi}}^{(2)}$. Afterwards, we can use equations (140), (153) and (146) in order to march $\hat{\Theta}^{(1)}$, $\hat{\hat{\Theta}}^{(2)}$ and η in time and the process can be repeated.

7 Acknowledgments

I would like to thank all fellows and staff for a great summer. Special thanks to Joe Keller and Raffaele Ferrari for supporting this work. Another special thanks to my advisor, Yehuda Agnon, for letting me know of this program and so much more. Finally, super special thanks deluxe to the upstairs office fellows (TLZ) for making this summer much more than the usual great GFD summers. May the force be with you.

References

[1] L. Rayleigh, Phi. Mag. **5**, (1876).

Two-dimensional Vortex Shedding From a Corner

David Vener

1 Introduction

The understanding of the unsteady separation of high-Reynolds number flow past a pointed edge is of interest to several fluid-mechanical problems, including flow past an aircraft wing or flow past a coastline. Much work has been done on the special case of flow past a semi-infinite line, notably by Rott [1], but much of that work has not been extended to less sharp points.

In this paper, we consider a two-dimensional approximation of high Reynolds number vortex shedding from a corner initially at rest in a motionless, unbounded, incompressible fluid. If the corner is sufficiently sharp, that is if it has exterior angle greater than π, vorticity must be shed when the fluid begins to move in order to maintain regularity at the edge. We consider a flow with a vortex sheet emanating from the corner at $t = 0$. As it moves away from the corner, the vortex sheet rolls up at its end producing an effect that is approximately that of a concentrated vortex. Therefore, we model the shed vorticity by a single point vortex.

The vortex created at $t = 0$ moves in the flow created by itself, it's image vortex, and the forcing flow. The magnitude of the vortex may not decrease because this would represent the unraveling of the sheet. Therefore, if it reaches some maximum at $t = t'$, a new vortex must be created in order to maintain regularity at the corner. The magnitude of the first vortex will remain constant while the the magnitude of the second vortex increases and both move in the flow created by the forcing and both vortices and their images. This process could continue ad infinitum each time the most recently created vortex reaches a maximum.

2 Mathematical Formulation

Let us consider a body with a (not necessarily bounded) boundary $\mathcal{C}$ in the physical plane, which we shall denote as the z-plane with $z = x+iy$ with $x, y \in \text{Re}$. Now suppose $F : z \to \zeta$ is a conformal mapping from the exterior of $\mathcal{C}$ in the physical plane to the plane $\text{Im}(\zeta) > 0$.

We create this image plane because we expect to be able to satisfy the boundary condition at the body more easily in this plane than in the physical plane. In the physical plane, we require that the velocity be parallel to the boundary. In the mapped plane, this is equivalent to

$$\text{Im}\Phi' \Big|_{\text{Im}\zeta=0} = 0. \tag{1}$$

Since a point vortex has a logarithmic singularity, we require

$$\Phi \sim \frac{\Gamma i}{2\pi} \log(\zeta - \zeta_0) + const. \tag{2}$$

near $\zeta = \zeta_0$, where Γ is the strength of the vortex. To satisfy the boundary condition, we place an "image vortex" with strength $-\Gamma$ at position $\zeta = \overline{\zeta_0}$ within the body in the lower half plane. The potential of the two vortices is

$$\Phi(\zeta) = \frac{i\Gamma}{2\pi} \left[\log(\zeta - \zeta_0) - \log(\zeta - \overline{\zeta_0}) \right] , \tag{3}$$

which satisfies the boundary condition (1).

To calculate the potential in the ζ-plane due to n vortices with strengths $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ at positions $\zeta_1, \zeta_2, \ldots, \zeta_n$ respectively, we need only superimpose the potentials due to the vortices and their images. Therefore, the potential including n vortices and the forcing flow, $U(t)$, is

$$\Phi(\zeta) = U(t)\,\zeta + \frac{i}{2\pi} \sum_{j=1}^{n} \Gamma_n \left[\log(\zeta - \zeta_j) - \log(\zeta - \overline{\zeta_j}) \right] . \tag{4}$$

Therefore, since $\Phi(\zeta)$ is analytic in the upper half plane and F is a conformal map, $\Phi(F^{-1}(z))$ is the complex potential in the physical plane which is created by n vortices at $z_1 = F(\zeta), z_2 = F(\zeta), \ldots, z_n = F(\zeta)$, their images, and a uniform flow. Thus, we can write the complex velocity in the physical plane, in terms of the variables in the ζ-plane. We have

$$\frac{d\Phi}{dz} = \frac{d\zeta}{dz}\Phi'(\zeta) \tag{5}$$

$$= \frac{1}{F'(\zeta)}\Phi'(\zeta). \tag{6}$$

Without loss of generality, we can choose the center of coordinates of the z-plane and F so that the corner in the physical plane is at $z = 0$ and so that $F(0) = 0$. Therefore, if $F'(0) = 0$, the only way to make the velocity at the origin nonsingular is to force $\Phi'(0) = 0$. This condition is known as the Kutta condition and can be expressed as

$$U(t) + \frac{i}{2\pi} \sum_{j=1}^{n} \Gamma_j \left[\frac{1}{\overline{\zeta_j}} - \frac{1}{\zeta_j} \right] = 0. \tag{7}$$

As each new vortex is created, n increases. Note that this condition only depends on variables defined in the ζ-plane.

We now derive the equation of motion for the vortices in the flow. Recall, that the velocity near a vortex has a singular part due to its strength and a nonsingular part due to the flow and all of the other vortices and images. Saffman [2] shows that the balance of the pressure force on a small circle around the vortex and the change of momentum through the circle requires that the vortex move with the nonsingular part of the flow. In the physical plane that means

$$\frac{d\overline{z_j}}{dt} = \lim_{z \to z_j} \frac{d}{dz} \left[\Phi \left(F^{-1}(z) \right) - \frac{i\Gamma_j}{2\pi} \log(z - z_j) \right]$$

$$= \lim_{\zeta \to \zeta_j} \frac{1}{F'(\zeta)} \frac{d}{d\zeta} \left[\Phi(\zeta) - \frac{i\Gamma_j}{2\pi} \log(F(\zeta) - F(\zeta_j)) \right]$$

Removing the terms which are not singular at $\zeta = \zeta_j$ and combining the singular logarithms, we have

$$\frac{d\overline{z_j}}{dt} = \frac{1}{F'(\zeta_j)} \left\{ U(t) + \frac{i}{2\pi} \sum_{k \neq j}^{n} \frac{\Gamma_k}{\zeta_j - \zeta_k} - \frac{\Gamma_k}{\zeta_j - \overline{\zeta_l}} - \frac{i\Gamma_j}{2\pi \left(\zeta_j - \overline{\zeta_j}\right)} - \frac{i\Gamma_j}{2\pi} \lim_{\zeta \to \zeta_j} \frac{d}{d\zeta} \log \frac{F(\zeta) - F(\zeta_j)}{\zeta - \zeta_j} \right\} \tag{8}$$

Finally, we expand $F(\zeta)$ in a Taylor series around ζ_j to find that

$$\begin{aligned} \log \frac{F(\zeta) - F(\zeta_j)}{\zeta - \zeta_j} &= \log \left[F'(\zeta_j) \left[1 + \frac{F''(\zeta_j)}{2F'(\zeta_j)} (\zeta - \zeta_j) + o\left((\zeta - \zeta_j)^2\right) \right] \right] \\ &= \log F'(\zeta_j) + \frac{F''(\zeta_j)}{2F'(\zeta_j)} (\zeta - \zeta_j) + o\left((\zeta - \zeta_j)^2\right) \end{aligned} \tag{9}$$

Substituting equation (9) into equation (8) reveals that

$$\frac{d\overline{z_j}}{dt} = \frac{1}{F'(\zeta_j)} \left\{ U(t) + \frac{i}{2\pi} \sum_{k \neq j}^{n} \frac{\Gamma_k}{\zeta_j - \zeta_k} - \frac{\Gamma_k}{\zeta_j - \overline{\zeta_l}} - \frac{i\Gamma_j}{2\pi \left(\zeta_j - \overline{\zeta_j}\right)} - \frac{i\Gamma_j}{4\pi} \frac{F''(\zeta_j)}{F'(\zeta_j)} \right\}. \tag{10}$$

The last term in equation (10) is known as the Routh correction and takes into account the self advection of the vortex. For convenience, we will study the evolution in the ζ-plane, so we convert equation (10) by conjugating both sides and multiplying on both sides by $d\zeta_j/dz_j = 1/F'(\zeta_j)$ to find that for each $1 \leq j \leq n$

$$\frac{d\zeta_j}{dt} = \frac{1}{|F'(\zeta_j)|^2} \left\{ U(t) - \frac{i}{2\pi} \sum_{k \neq j}^{n} \frac{\Gamma_k}{\overline{\zeta_j} - \overline{\zeta_k}} - \frac{\Gamma_k}{\overline{\zeta_j} - \zeta_k} - \frac{i\Gamma_j}{2\pi \left(\zeta_j - \overline{\zeta_j}\right)} + \frac{i\Gamma_j}{4\pi} \frac{\overline{F''(\zeta_j)}}{\overline{F'(\zeta_j)}} \right\}. \tag{11}$$

For completeness, we mention here that some authors have used equations other than equation (11) to study the shedding of vortices in two-dimensional flows. Brown and Michael [3] include a correction term on the left-hand side of equation (11) which is proportional to $d\Gamma_n/dt$ to balance the force on a hypothetical line of vorticity stretching from the corner to the vortex, which feeds the vortex allowing it to grow in strength. Howe [4] also considers a line of vorticity, but he further requires that the correction term account for balancing the torque on the sheet. We will discuss Brown and Michael's equation briefly in section 3, but will not consider Howe's equation because it is significantly more complicated and even according to his own results only adds a very small correction.

3 The Evolution of the First Vortex

As discussed in Section 2, at $t = 0$, i.e. just as the fluid starts moving, we expect a vortex to be shed from the corner in the physical plane. Since there is only one vortex in the fluid initially, we need only consider equation (11) for $j = n = 1$ where Γ_1 allows the Kutta condition, i.e. equation (7), to be satisfied. Solving equation (7) for Γ_1 yields

$$\Gamma_1 = 2\pi i \frac{|\zeta_1|^2}{\zeta_1 - \overline{\zeta_1}} U(t). \tag{12}$$

Therefore, we have

$$\begin{aligned}\frac{d\zeta_1}{dt} &= \frac{1}{|F'(\zeta_1)|^2}\left[U(t) - \frac{i\Gamma_1}{2\pi\ \ \zeta_1 - \overline{\zeta_1}} + \frac{i\Gamma_1}{4\pi}\frac{\overline{F''(\zeta_1)}}{\overline{F'(\zeta_1)}}\right] \\ &= \frac{U(t)}{|F'(\zeta_1)|^2}\left[1 - \frac{|\zeta_1|^2}{4\left(\mathrm{Im}\zeta_1\right)^2} + i\frac{\overline{F''(\zeta_1)}}{\overline{F'(\zeta_1)}}\frac{|\zeta_1|^2}{4\mathrm{Im}\zeta_1}\right], \end{aligned} \tag{13}$$

with the initial condition that $\zeta_1(0) = 0$.

3.1 An Exact Solution for the Infinite Wedge

To proceed, we must choose the boundary in the physical plane and calculate the conformal map F. An infinite wedge with its tip at $z = 0$ is of particular interest because in the region very near the tip, any corner is approximated by the infinite wedge. Therefore, let us choose the boundary to be an infinite wedge with exterior angle $\alpha\pi$ that is bisected by the imaginary axis. The corresponding conformal map is therefore given by

$$F(\zeta) = e^{-i(\alpha-1)\frac{\pi}{2}}\zeta^{\alpha}, \tag{14}$$

so that equation (13) becomes

$$\frac{d\zeta_1}{dt} = \frac{U(t)}{\alpha^2\left|\zeta_1\right|^{2(\alpha-1)}}\left[1 - \frac{|\zeta_1|^2}{4(\mathrm{Im}\zeta_1)^2} + i\frac{(\alpha-1)\,\zeta_1}{4\mathrm{Im}\zeta_1}\right]. \tag{15}$$

Writing $\zeta = \xi + i\eta$ with ξ, η real, we separate real and imaginary parts in equation (15) to find coupled equations for $d\xi_1/dt$ and $d\eta_1/dt$. That is

$$\frac{d\xi_1}{dt} = \frac{U(t)}{\alpha^2\ \ \xi_1^2 + \eta_1^2\ ^{\alpha-1}}\ \ 1 - \frac{\alpha}{4} - \frac{\xi_1^2}{4\eta_1^2}\ \ , \tag{16}$$

and

$$\frac{d\eta_1}{dt} = \frac{U(t)}{\alpha^2\ \ \xi_1^2 + \eta_1^2\ ^{\alpha-1}}\ \ \frac{\alpha-1}{4}\frac{\xi_1}{\eta_1}\ \ . \tag{17}$$

To solve equations (16) and (17), we note that the only explicit time dependence occurs in a prefactor that multiplies the right side of both equations. Therefore, by dividing equation (16) by equation (17) we are left with a differential equation for ξ_1 as a function of η_1. Solving this equation with the initial condition $\xi(\eta = 0) = 0$ yields gives us

$$\xi_1 = \pm\sqrt{\frac{4-\alpha}{\alpha}}\eta_1. \tag{18}$$

Now, we can solve for $\eta_1(t)$ in equation (17) by substituting equation (18) for ξ_1 recalling that $\eta_1 \geq 0$. This yields an exact solution for η_1, which in turn gives us an exact solution for ζ_1. We find

$$\zeta_1(t) = \frac{(\alpha-1)(2\alpha-1)}{4\alpha^{5/2}}\left|\int_0^t U(\tau)\,d\tau\right|^{\frac{1}{2\alpha-1}}\ \ \frac{\sqrt{\alpha}}{2}i + sign\ \ \int_0^t U(\tau)\,d\tau\ \ \frac{\sqrt{4-\alpha}}{2}\ \ . \tag{19}$$

This indicates that the first vortex moves away from the origin along a ray in the ζ-plane.

3.2 Preparing to Shed the Second Vortex

As discussed in Section (2), we expect the vortex to continue along its path as determined by equation (19) until such time that the magnitude of its strength reaches a maximum. From equation (12), we compute

$$\Gamma_1(t) = \frac{2\pi U(t)}{\sqrt{\alpha}} \quad \frac{(\alpha-1)(2\alpha-1)}{4\alpha^{5/2}} \left| \int_0^t U(\tau)\, d\tau \right|^{\frac{1}{2\alpha-1}}. \tag{20}$$

Without loss of generality, we can take U to be nonnegative for small t since replacing U with $-U$ and ξ_1 with $-\xi_1$ leaves equations (16) and (17) unchanged. With this assumption, upon differentiating equation (20) by t, we see that Γ_1 reaches a local extremum at $t = \hat{t}$ if and only if

$$U' \;\; \hat{t} \;\; \int_0^{\hat{t}} U(\tau)\, d\tau + \frac{U^2 \;\; \hat{t}}{2\alpha - 1} = 0. \tag{21}$$

3.3 An Exact Solution with the Brown and Michael Model

As discussed in Section 2, Brown and Michael [3] added a correction term to equation (10) to account for an unbalanced force on a vortex sheet leading up to the vortex coming out of the origin. For the first vortex, the corrected equation is

$$\frac{d\overline{z_1}}{dt} = -\overline{z_1}\frac{1}{\Gamma_1}\frac{d\Gamma_1}{dt} + \frac{1}{F'(\zeta_j)}\left\{ U(t) - \frac{i\Gamma_j}{2\pi \;\; \zeta_j - \overline{\zeta_j}} - \frac{i\Gamma_j}{4\pi}\frac{F''(\zeta_j)}{F'(\zeta_j)} \right\}. \tag{22}$$

Cortelezzi [5] gives an exact solution for the infinite wedge in the case that $\alpha = 2$. However, we can solve equation (22) for any value of α, $1 < \alpha \leq 2$. In terms of ζ_1 equation (22) can be rewritten

$$\alpha\overline{\zeta_1}^{\alpha-1}\dot{\overline{\zeta_1}} = -\overline{\zeta_1}^{\alpha} \quad \frac{\dot{U}}{U} + \frac{\dot{\overline{\zeta_1}}}{\overline{\zeta_1}} + \frac{\dot{\zeta_1}}{\zeta_1} - \frac{\dot{\zeta_1} - \dot{\overline{\zeta_1}}}{\zeta_1 - \overline{\zeta_1}} \Bigg) + \frac{U(t)}{\alpha\zeta_1^{\alpha-1}} \quad 1 + \frac{|\zeta_1|^2}{\zeta_1 - \overline{\zeta_1}^{\,2}} + \frac{\alpha-1}{2}\frac{\overline{\zeta_1}}{\zeta_1 - \overline{\zeta_1}} \Bigg). \tag{23}$$

Changing variables by setting $\zeta = \rho e^{i\theta}$ reduces equation (23)to equations for $\dot{\rho}$ and $\dot{\theta}$. They are

$$\dot{\theta} = \frac{\alpha - 4\sin^2\theta}{4\alpha^2 \sin\theta}\frac{U(t)}{\rho^{2\alpha-1}}, \tag{24}$$

and

$$\dot{\rho} = -\frac{\rho}{\alpha+1}\frac{\dot{U}}{U} + \frac{\alpha-1}{\alpha^2(\alpha+1)}\frac{U\cos\theta}{\rho^{2\alpha-2}}. \tag{25}$$

We now solve equations (24) and (25) using the following change of variables:

$$\lambda = U^{(2\alpha-1)/(\alpha+1)}\rho^{2\alpha-1}, \quad \mu = \cos\theta, \quad \tilde{t} = \int_0^t U^{3\alpha/(\alpha+1)}(\tau)\, d\tau. \tag{26}$$

After some algebra, we have

$$\frac{d\lambda}{d\tilde{t}} = \frac{(2\alpha-1)(\alpha-1)}{\alpha^2(\alpha+1)}\mu, \quad \frac{d\mu}{d\tilde{t}} = \frac{4-\alpha-4\mu^2}{4\alpha^2\lambda} \tag{27}$$

with the initial conditions that $\lambda(0) = 0$ and $\mu(0) = \mu_0$. Combining equations (27) to form a single differential equation for $u = \lambda^2$ gives

$$\frac{d^2u}{d\tilde{t}^2} = \frac{(2\alpha-1)(\alpha-1)(4-\alpha)}{2\alpha^4(\alpha+1)} - \frac{\alpha(2-\alpha)}{(2\alpha-1)(\alpha-1)}\frac{1}{u}\left(\frac{du}{d\tilde{t}}\right)^2. \tag{28}$$

To solve this equation, we set $du/d\tilde{t} = f(u)$ since there is no explicit time dependence. After some manipulation, we find that

$$\lambda\ \tilde{t}\ = \pm\frac{(2\alpha-1)(\alpha-1)}{2\alpha^2(\alpha+1)}\sqrt{4-\alpha}\tilde{t} \tag{29}$$

and that

$$\mu = \pm\frac{\sqrt{4-\alpha}}{2}. \tag{30}$$

Putting all of this together, we find that the exact solution for the first vortex with the Brown and Michael correction is

$$\zeta_1 = \left(\frac{(2\alpha-1)(\alpha-1)\sqrt{4-\alpha}}{2\alpha^2(\alpha+1)}\right)^{\frac{1}{2\alpha-1}}\left(\frac{1}{U(t)}\right)^{\frac{1}{\alpha+1}}\left(\int_0^t U(\tau)^{3\alpha/(\alpha+1)}\,d\tau\right)^{\frac{1}{2\alpha-1}}\left(\pm\frac{\sqrt{4-\alpha}}{2}+i\frac{\sqrt{\alpha}}{2}\right). \tag{31}$$

Therefore,

$$\Gamma_1 = \frac{2\pi}{\sqrt{\alpha}}\left(\frac{(2\alpha-1)(\alpha-1)\sqrt{4-\alpha}}{2\alpha^2(\alpha+1)}\right)^{\frac{1}{2\alpha-1}}U^{\frac{\alpha(2\alpha-1)}{\alpha+1}}\int_0^t (U(\tau))^{\frac{3\alpha}{\alpha+1}}\,d\tau, \tag{32}$$

which implies that the first vortex increases in strength until $t = t_1$ satisfies

$$\frac{\alpha(2\alpha-1)}{\alpha+1}U'(t_1)\int_0^{t_1}(U(\tau))^{\frac{3\alpha}{\alpha+1}}\,d\tau + U(t_1)^{\frac{4\alpha+1}{\alpha+1}} = 0. \tag{33}$$

This agrees with the solution found by Cortelezzi for $\alpha = 2$. In addition, equation (31) agrees with our solution without the vortex sheet, equation (19), in its angle of departure.

4 The Second Vortex for the Infinite Wedge

If for some $t_1 > 0$, equation (21) is satisfied, a second vortex will be shed from the corner of the infinite wedge, and the strength of the first vortex will be fixed for all times $t \geq t_1$. For example, if $U = \sin t$, $t_1 = \cos^{-1}(-1/2\alpha)$. The system of equations given by (11) with $n = 2$ is

$$\frac{d\zeta_1}{dt} = \frac{1}{\alpha^2|\zeta_1|^{2\alpha-2}}\left\{U(t) - \frac{i}{2\pi}\left(\frac{\Gamma_2}{\overline{\zeta_1}-\overline{\zeta_2}} - \frac{\Gamma_2}{\overline{\zeta_1}-\zeta_2}\right) - \frac{i\Gamma_1}{2\pi\left(\zeta_1-\overline{\zeta_1}\right)} + (\alpha-1)\frac{i\Gamma_1}{4\pi\overline{\zeta_1}}\right\}, \tag{34}$$

and

$$\frac{d\zeta_2}{dt} = \frac{1}{\alpha^2|\zeta_2|^{2\alpha-2}}\left\{U(t) - \frac{i}{2\pi}\left(\frac{\Gamma_1}{\overline{\zeta_2}-\overline{\zeta_1}} - \frac{\Gamma_1}{\overline{\zeta_2}-\zeta_1}\right) - \frac{i\Gamma_2}{2\pi\left(\zeta_2-\overline{\zeta_2}\right)} + (\alpha-1)\frac{i\Gamma_2}{4\pi\overline{\zeta_2}}\right\}, \tag{35}$$

with the initial conditions that ζ_1 is continuous at t_1 and that $\zeta_2(t_1) = 0$.

From the Kutta condition, we can solve for Γ_2 yielding

$$\Gamma_2 = 2\pi i \frac{|\zeta_2|^2}{\zeta_2 - \overline{\zeta_2}} \quad U(t) + \frac{i\Gamma_1}{2\pi} \quad \frac{1}{\overline{\zeta_1}} - \frac{1}{\zeta_1} \quad . \tag{36}$$

Plugging equation (36) into equation (35), we find

$$\frac{d\zeta_2}{dt} = \frac{1}{\alpha^2 |\zeta_2|^{2\alpha-2}} \left\{ U(t) - \frac{i}{2\pi} \quad \frac{\Gamma_1}{\overline{\zeta_2} - \overline{\zeta_1}} - \frac{\Gamma_1}{\overline{\zeta_2} - \zeta_1} \quad + \quad U + \frac{i\Gamma_1}{2\pi} \frac{\zeta_1 - \overline{\zeta_1}}{|\zeta_1|^2} \quad \left[-\frac{|\zeta_1|^2}{4(\mathrm{Im}\zeta_2)^2} + i(\alpha - 1)\frac{\zeta_2}{4\mathrm{Im}\zeta_2} \right] \right\} \tag{37}$$

Since $0 < |\zeta_2| \ll |\zeta_1(t)|$ for $t - t_1 \ll 1$, we can expand the second term inside the braces in equation (37). Defining

$$K_1(t) = U + \frac{i\Gamma_1}{2\pi} \frac{\zeta_1 - \overline{\zeta_1}}{|\zeta_1|^2}, \tag{38}$$

and

$$\epsilon_2 = \frac{1}{\alpha^2 |\zeta_2|^{2\alpha-2}}, \tag{39}$$

we are able to write

$$\frac{d\zeta_2}{dt} = \epsilon_2 \left\{ K_1(t) \left[1 - \frac{|\zeta_2|^2}{4(\mathrm{Im}\zeta_2)^2} + i\frac{(\alpha-1)\zeta_2}{4\mathrm{Im}\zeta_2} \right] - 2\mathrm{Re}(g_1)\overline{\zeta_2} - 2\mathrm{Re}(h_1)\overline{\zeta_2}^2 + o(\zeta_2^3) \right\}, \tag{40}$$

where

$$g_1 = \tfrac{i\Gamma_1}{2\pi\zeta_1^2}, \quad h_1 = \tfrac{i\Gamma_1}{2\pi\zeta_1^3}. \tag{41}$$

4.1 The Size of the Terms in Equation (40)

Recall that we defined

$$K_1(t) = U + \frac{i\Gamma_1}{2\pi} \frac{\zeta_1 - \overline{\zeta_1}}{|\zeta_1|^2} \tag{42}$$

for all times t. From the Kutta condition, $K_1 \equiv 0$ for all $t < t_1$. Therefore,

$$\begin{aligned} \frac{d}{dt}K_1 &= U' + \frac{i\Gamma_1}{2\pi}\frac{d}{dt} \quad \frac{\zeta_1 - \overline{\zeta_1}}{|\zeta_1|^2} \quad + \frac{i\Gamma_1'}{2\pi}\frac{\zeta_1 - \overline{\zeta_1}}{|\zeta_1|^2} \\ &= 0 \end{aligned} \tag{43}$$

for all $t < t_1$. However, t_1 is defined to be the location of the first local extremum of Γ_1 so that $\Gamma_1'(t_1) = 0$. Therefore,

$$\begin{aligned} \lim_{t \to t_1^+} \frac{dK_1}{dt}(t) &= U'(t_1) + \frac{i\Gamma_1(t_1)}{2\pi}\frac{d}{dt} \quad \frac{\zeta_1 - \overline{\zeta_1}}{|\zeta_1|^2} \quad \Big|_{t=t_1} \\ &= 0, \end{aligned} \tag{44}$$

since U and ζ_1 are continuous functions at $t = t_1$. This means that the Taylor expansion of K_1 for $t > t_1$ is of the form

$$K_1(t) = \frac{k_1}{2}(t - t_1)^2 + o\ (t - t_1)^3 \quad , \tag{45}$$

where

$$k_1 = \lim_{t \to t_1^+} \frac{d^2 K_1}{dt^2}(t) \tag{46}$$

is, in general, a nonzero constant.

Now if we suppose that $\zeta_2 \sim (t - t_1)^\beta$, equating exponents of the leading terms on both sides of equation (40), we find that

$$\beta - 1 = \min(2, \beta) - 2\beta(\alpha - 1) \tag{47}$$

so that

$$\beta = \begin{cases} \frac{1}{2\alpha - 2} & \alpha \geq \frac{5}{4} \\ \frac{3}{2\alpha - 1} & \alpha \leq \frac{5}{4}. \end{cases} \tag{48}$$

4.2 Case 1: $\alpha > 5/4$

From equation (48), we see that if $\alpha > 5/4$, then to leading order in $(t - t_1)$, equation (40) is

$$\frac{d\zeta_2}{dt} = -\frac{2\mathrm{Re}(g_1(t_1))}{\alpha^2 |\zeta_2|^{2\alpha - 2}} \overline{\zeta_2}. \tag{49}$$

Therefore, setting $\zeta_2 \sim a(t - t_1)^{1/(2\alpha - 2)}$, we find an equation for a:

$$\frac{a}{2\alpha - 2} = -2\frac{\overline{a}}{\alpha^2 |a|^{2\alpha - 2}} \mathrm{Re}(g_1(t_1)). \tag{50}$$

Solving for a, we find

$$a = \frac{4\alpha - 4}{\alpha^2} |\mathrm{Re}(g_1(t_1))|^{2\alpha - 2} \sqrt{sign(-\mathrm{Re}(g_1(t_1)))}. \tag{51}$$

Therefore, if $\mathrm{Re}(g_1) > 0$,

$$\zeta_2 \sim i \frac{4\alpha - 4}{\alpha^2} |\mathrm{Re}(g_1(t_1))|^{2\alpha - 2} (t - t_1)^{1/(2\alpha - 2)} \tag{52}$$

and

$$\Gamma_2 \sim \pi \frac{k_1}{2} \frac{4\alpha - 4}{\alpha^2} |\mathrm{Re}(g_1(t_1))|^{2\alpha - 2} (t - t_1)^{2 + \frac{1}{2\alpha - 2}}. \tag{53}$$

However, if $\mathrm{Re}(g_1(t_1)) < 0$, we would have a contradiction because the vortex would be trying to stay on the wedge so that the boundary condition cannot be satisfied. Therefore, we must determine $sign(\mathrm{Re}(g_1))$. We calculate,

$$\begin{aligned} sign(\mathrm{Re}(g_1)) &= sign \frac{i\Gamma_1}{2\pi\zeta_1^2} \\ &= sign(\Gamma_1)\, sign\ \mathrm{Im}\ \zeta_1^2 \quad . \end{aligned} \tag{54}$$

We know that if $signU > 0$, then ζ_1 is in the first quadrant so that ζ_1^2 is in the upper half plane; however, ζ_1 is in the second quadrant if $signU < 0$, which implies that ζ_1^2 is in the lower half plane. Therefore, $sign\ \mathrm{Im}\ \zeta_1^2 = signU$, and

$$\begin{aligned} sign\left(\mathrm{Re}\left(g_1\right)\right) &= sign(\Gamma_1)sign(U) \\ &= \left(sign(U)\right)^2 > 0, \end{aligned} \tag{55}$$

which implies that the second vortex can always be released. Once the vortex moves away from the singular point at the origin, the motion of both vortices can be analyzed by numerically integrating equations (34) and (35).

4.3 Case 2: $\alpha < 5/4$

Setting $\zeta_2 = \xi_2 + i\eta_2$, we see that equation (40) is, to leading order in $(t - t_1)$,

$$\frac{d\zeta_2}{dt} = \frac{1}{\alpha^2 \left|\zeta_2\right|^{2\alpha-2}} \left\{ \frac{k_1}{2}\left(t-t_1\right)^2 \quad 1 - \frac{\alpha}{4} - \frac{\xi_2^2}{4\eta_2^2} + \left(\alpha+1\right) i\frac{\xi_2}{\eta_2} \quad \right\}. \tag{56}$$

This is equivalent to equation (15) with ζ_1 replaced with ζ_2 and $U\left(t\right)$ replaced with $k_2\left(t-t_1\right)^2/2$. Therefore,

$$\zeta_2 \sim \quad \frac{\left(\alpha-1\right)\left(2\alpha-1\right)}{12\alpha^{5/2}} \left|k_1\right| \left(t-t_1\right)^3 \quad ^{\frac{1}{2\alpha-1}} \quad \frac{\sqrt{\alpha}}{2} i + sign\left(k_1\right) \frac{\sqrt{4-\alpha}}{2} \quad , \tag{57}$$

and

$$\Gamma_2 \sim \frac{\pi k_1 \left(t-t_1\right)^2}{\sqrt{\alpha}} \quad \frac{\left(\alpha-1\right)\left(2\alpha-1\right)}{12\alpha^{5/2}} \left|k_1\right| \left(t-t_1\right)^3 \quad ^{\frac{1}{2\alpha-1}} \tag{58}$$

shortly after the second vortex is created. After it moves away from the origin, we can again use numerical integration to explore the dynamics of the system.

5 More Vortices for the Infinite Wedge

Now suppose that at some $t_2 > t_1$, the magnitude of the second vortex reaches a maximum. If we again assume that, once this maximum is reached, the strength of the second vortex is fixed, a third vortex must be shed from the origin to satisfy the Kutta condition. Furthermore, under certain conditions which will be discussed in this section, this process could repeat itself several (perhaps, infinitely many) more times. Therefore, for the rest of the section let us suppose that n vortices have been released from the origin and they are located at $z_1, z_2, \ldots, z_n$ respectively in the physical plane (corresponding to $\zeta_1, \zeta_2, \ldots, \zeta_n$ in the image plane). We further assume that each of the ζ_j are nonzero and that there exist times $0 < t_1 < t_2 < \cdots < t_n$ at which each of the n vortices have reached a maximum, respectively. We now consider the dynamics as the next vortex is released, i.e. for $t > t_n$.

Using equation (11) for a system of $n+1$ vortices and equation (14) we have for the newest vortex

$$\frac{d\zeta_{n+1}}{dt} = \epsilon_{n+1} \left\{ U\left(t\right) - \frac{i}{2\pi} \sum_{k=1}^{n} \quad \frac{\Gamma_k}{\overline{\zeta_{n+1}} - \overline{\zeta_k}} - \frac{\Gamma_k}{\overline{\zeta_{n+1}} - \zeta_k} \quad - \frac{i\Gamma_{n+1}}{2\pi \quad \zeta_{n+1} - \overline{\zeta_{n+1}}} + i\frac{\left(\alpha-1\right)\Gamma_{n+1}}{4\pi\overline{\zeta_{n+1}}} \right\}, \tag{59}$$

while for each j, $1 \leq j \leq n$,

$$\frac{d\zeta_j}{dt} = \epsilon_j \left\{ U(t) - \frac{i}{2\pi} \sum_{k \neq j}^{n+1} \frac{\Gamma_k}{\overline{\zeta_j} - \overline{\zeta_k}} - \frac{\Gamma_k}{\overline{\zeta_j} - \zeta_k} - \frac{i\Gamma_j}{2\pi \left(\zeta_j - \overline{\zeta_j}\right)} + i\frac{(\alpha-1)\Gamma_j}{4\pi\overline{\zeta_j}} \right\}, \tag{60}$$

where we have defined, in agreement with section 4,

$$\epsilon_l = \frac{1}{\alpha^2 |\zeta_l|^{2\alpha-2}} \tag{61}$$

for each l, $1 \leq l \leq n+1$. These equations must be solved with the initial conditions

$$\zeta_{n+1}(t_n) = 0, \quad \Gamma_{n+1}(t_n) = 0, \tag{62}$$

for the new vortex and that each of the other vortices must move in a continuous manner, holding their strength constant. From the Kutta Condition, i.e. equation (7) with $n+1$ vortices, we can solve for Γ_{n+1}, finding

$$\Gamma_{n+1} = 2\pi i \frac{|\zeta_{n+1}|^2}{\zeta_{n+1} - \overline{\zeta_{n+1}}} \left[U(t) + \sum_{k=1}^{n} \frac{i\Gamma_k}{2\pi} \left(\frac{1}{\overline{\zeta_k}} - \frac{1}{\zeta_k} \right) \right]. \tag{63}$$

We can substitute equation (63) into equation (59) and expand the resulting equation to leading order in ζ_{n+1} since, for $t - t_n \ll 1$, $\zeta_{n+1} \ll \zeta_j$, for all $1 \leq j \leq n$. The resulting equation of motion is

$$\frac{d\zeta_{n+1}}{dt} = \epsilon_{n+1} \left\{ K_n(t) \left[1 - \frac{|\zeta_{n+1}|^2}{4\left(\mathrm{Im}\zeta_{n+1}\right)^2} + i\frac{(\alpha-1)\zeta_{n+1}}{4\mathrm{Im}\zeta_{n+1}} \right] - 2\mathrm{Re}(g_n)\overline{\zeta_{n+1}} - 2\mathrm{Re}(h_n)\overline{\zeta_{n+1}}^2 + o(\zeta_{n+1}^3) \right\}, \tag{64}$$

with

$$K_n(t) = U(t) + \sum_{k=1}^{n} \frac{i\Gamma_k}{2\pi} \frac{\zeta_k - \overline{\zeta_k}}{|\zeta_k|^2}, \tag{65}$$

and

$$g_n = \textstyle\sum_{k=1}^{n} \frac{i\Gamma_k}{2\pi\zeta_k^2}, \quad h_n = \sum_{k=1}^{n} \frac{i\Gamma_k}{2\pi\zeta_k^3}. \tag{66}$$

Note that equation (64) has exactly the same form as equation (40). Furthermore, the same reasoning used in section 4.1 to show that $K_1 \sim (t - t_1)^2$ can be reapplied to show that

$$K_n(t) \sim \frac{k_n}{2}(t - t_n)^2 \tag{67}$$

for $t - t_n \ll 1$ with

$$k_n = U'(t_n) + \left. \frac{d}{dt} \sum_{k=1}^{n} \frac{i\Gamma_k(t_n)}{2\pi} \frac{\zeta_k - \overline{\zeta_k}}{|\zeta_k|^2} \right) \right|_{t=t_n}. \tag{68}$$

Supposing that $\zeta_{n+1} \sim (t - t_n)^\beta$, equating the exponents of the leading terms in equation (64) reveals that

$$\beta = \begin{cases} \frac{1}{2\alpha-2} & \alpha \geq \frac{5}{4} \\ \frac{3}{2\alpha-1} & \alpha \leq \frac{5}{4} \end{cases} \tag{69}$$

as before.

5.1 Revisiting $\alpha > 5/4$

Following section 4.2, we find that setting $\zeta_{n+1} \sim a\,(t - t_n)^{1/(2\alpha-2)}$ and substituting in equation (64) gives us an equation for a which can be solved to give

$$a = \frac{4\alpha - 4}{\alpha^2} \left|\mathrm{Re}(g_n\,(t_n))\right|^{2\alpha-2} \sqrt{sign(-\mathrm{Re}(g_n(t_n)))} \tag{70}$$

which still leaves us with a contradiction if $\mathrm{Re}(g_n(t_n)) < 0$. However, with more than two vortices it is actually possible for this contradiction to manifest itself. In fact, for the infinite line ($\alpha = 2$), considered by Cortelezzi and others, and $U = \sin t$, this contradiction arises at t_2, so that the third vortex cannot be shed. If we choose

$$U\,(t) = \begin{cases} \frac{2-2\pi\tau}{\tau} t + \frac{2\pi\tau-1}{\tau^2} t^2, & t \le \tau \\ 1 + \frac{1}{10}\cos(20\pi t), & t > \tau, \end{cases} \tag{71}$$

with $\tau = .075$ following Cortelezzi [5], $g_n\,(t_n)$ is always positive so that vortices can be shed indefinitely.

5.2 Vorticity Dipoles

To handle the contradiction just discussed, we must consider what happens when vorticity is held near the boundary. The contradiction arises because the newly shed vortex and its image are at the same point in space when the vortex is on the boundary. Rott [1] suggests that one possibility would be to consider a boundary layer of vorticity around the body. A simpler idea is to treat the vortex and its image together as a vorticity dipole with strength $\vec{D}$ chosen to be parallel to the boundary. However, this approximation has a few complications. First, the strength of the dipole is not the same in the physical and mapped planes so that the shedding condition either cannot be treated as only dependent on variables in the mapped plane or does not correspond to the strength of the dipole reaching a maximum. Second, there is some ambiguity as to which direction along the boundary the dipole travels; that is, two possible trajectories appear to satisfy the Kutta condition at the origin and the equations of motion. These complications have not been worked out yet and will be the subject of future work.

6 Acknowledgments

I would like to thank Joe Keller and Stefan Llewellyn-Smith for supervising my work at Woods Hole and for many useful discussions. The other fellows, especially my officemates on TLZ, also deserve much thanks for making my summer so enjoyable. Finally, I would like to thank all of the participants in the program who make WHOI such an exciting place to live and work for the summer.

References

[1] N. Rott, "Diffraction of a weak shock with vortex generation," J. Fluid Mech. **1**, 111 (1956).

[2] P. Saffman, *Vortex Dynamics* (Cambridge University Press, Cambridge, 1992).

[3] C. Brown and W. Michael, "Effect of leading-edge separation on the lift of a delta wing," J. Aerosp. Sci. **21**, 690 (1954).

[4] M. Howe, "Emendation of the brown and michael equation," J. Fluid Mech. **329**, 89 (1996).

[5] L. Cortelezzi, "On the unsteady separated flow past a semi-infinite plate," Phys. Fluids **7**, 526 (1995).

Laboratory experiments on the effect of baroclinic eddies on a dense plume in a rotating stratified fluid

Danielle Wain

1 Introduction

Eighteen degree water (EDW) is a subtropical mode water of the North Atlantic. Mode water is a layer of water with vertically constant temperature and salinity. It often forms near strong baroclinic fronts, which provide a mechanical forcing for its creation and properties, in regions of high air-sea interaction, a buoyant forcing. Figure 1 shows the region of high wintertime net heat flux, which corresponds with the location of the EDW. This region of large heat flux is located near the Gulf Stream, a strong baroclinic front. Figure 2 shows the mode water. The left panel shows the temperature and velocity in this region. The thick layer of constant temperature can be seen here as well as the front from the Gulf Stream. The center panel shows the constant salinity layer and the isopycnals, which clearly show the mode water. The right panel shows the historical formation of the EDW predicted from air-sea fluxes, which in recent times has ranged between 15 and 25 Sv [1]. But Kwon and Riser inferred that 4 Sv is injected into the subtropical gyre[2]. The discrepancy between the estimated and the observed values is currently under investigation. The goal of these experiments is to investigate if lateral eddy fluxes by the eddies created by the baroclinic front of the Gulf Stream could explain the discrepancy. In the current experiments, a baroclinic front was created via upwelling and a dense line plume was injected at the surface to simulate the processes under which the EDW is formed. From hereafter, a line plume is defined as a plume that is driven by a buoyancy flux per unit length on a line whose extent in one horizontal direction is much larger than in the other horizontal direction, whereas a point plume is emitted from a single point.

The EDW formation process was simulated as a line plume in a rotating linearly stratified fluid. In a non-rotating flow, Rouse et al. examined the case of a line plume in a homogeneous fluid[4]. Morton et al. developed a theory for convection for both continuous and instantaneous point sources for an arbitrary stratification (including the homogeneous case), with a detailed analysis for the linearly stratified case[5]. Wright and Wallace experimentally examined the behavior of a line plume in a linearly stratified fluid[6]. Other previous experimental work on plumes in stratified fluids includes that of Noh et al. and Ching et al., who examined respectively a thermal and a negatively buoyany turbulent line plume descending through a two-layer fluid with a sharp density interface[7][8]. Laboratory experiments on the effect of rotation on a bouyant line plume in a homogeneous fluid were conducted by Fernando and Ching[9]. Lavelle and Smith studied this effect numerically

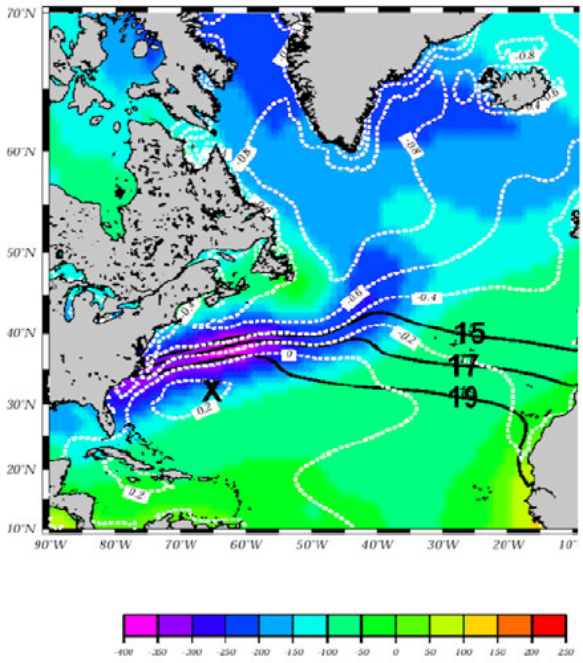


Figure 1: Wintertime net heat flux over the North Atlantic (Steve Riser, from John Marshall, pers. comm.).

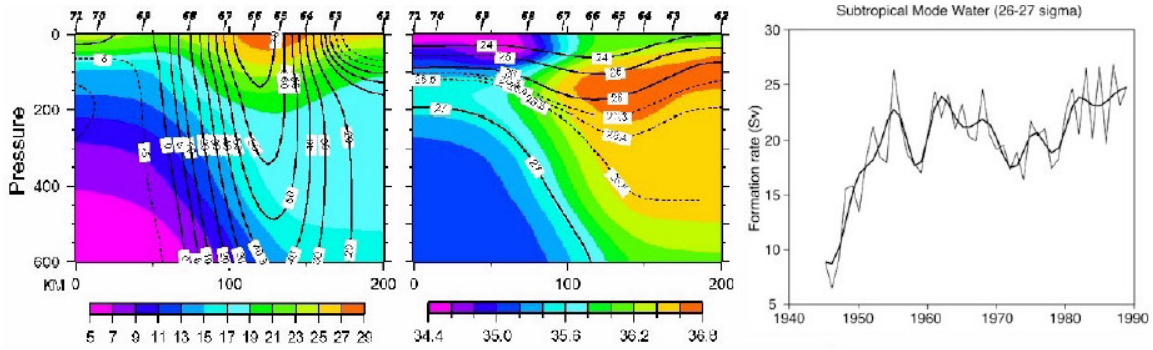


Figure 2: Meridional sections across the Gulf Stream of temperature and velocity(left) and salinity and potential density (center) (from Joyce[3]). The right panel shows the formation rate of the subtropical mode water estimated from air-sea fluxes.

using a three-dimensional nonhydrostatic model[10]. Fernando et al. performed laboratory experiments on the development of a dense turbulent point plume in a rotating homogeneous fluid[11]. More recently, Bush and Woods experimentally investigated the vortices produced by a buoyant line plume in both a homogeneous and a linearly stratified flow[12]. The rotating stratified case for a point plume has been studied by Helfrich and Battisti, who conducted laboratory experiments studying the effect of rotation on a hydrothermal plume[13], and Speer and Marshall, who conducted numerical experiments on the convection produced by a buoyant point source in a linearly stratified fluid[14].

The previous work on plumes most relevant to the current study is the experiments of Bush and Woods on a line plume in a rotating linearly stratified fluid. From their experiments, they found that the height of intrusion of the plume was

$$Z_n = (3.0 \pm 1.0)\frac{(B/L)^{1/3}}{N}, \tag{1}$$

where B is the total buoyancy flux, L is the length of the source, and N is the buoyancy frequency (which is constant in a linearly stratified fluid)[12]. The coefficient is similar to the value of 2.5 found by Wright and Wallace[6]. The reduced gravity of the intrusion in their experiments was

$$g_i' = g\frac{\rho_i - \rho_b}{\rho_b} \tag{2}$$

where ρ_i is the density of the intrusion and ρ_b is the density at the bottom of their tank (where the injection occurred in their experiments). They did not measure the density of the intrusion but predicted that it scaled as

$$g'(Z_n) \sim (B/L)^{1/3}N. \tag{3}$$

Manipulation of their scaling for the size of the eddies formed by the plume gives an eddy radius that scales as

$$R_{plume} \sim NZ_n/f,$$

where f is the Coriolis parameter.

Others have performed experiments investigating the eddies formed from fronts created by upwelling. In these experiments the eddies are created by instabilities at the front. For example, Narimousa and Maxworthy studied a two layer fluid where upwelling was forced by a disk spinning on the surface of the fluid[15]. In these experiments they found the velocity on the surface was

$$V = 0.74r\Delta\Omega, \tag{4}$$

where r is the radius of the disk and $\Delta\Omega$ is the rotation rate of the disk (with respect to the fluid). In other experiments, eddies were observed that scaled with the Rossby radius of deformation:

$$R = NH/f, \tag{5}$$

where H is the depth of the fluid[16].

The current experiments examined the effect of these eddies created by upwelling on the dense plume injected into a linearly stratified rotating fluid.

2 Experiment

The experiments were conducted in circular tank with a diameter of 115 cm positioned in the center of a 2 m rotating table. A disk with a radius r of 13 cm was suspended over the center of the tank at a height of 31 cm above the bottom. The linear salt stratification was created using the double-bucket method. The tank was filled while rotating until the disk was resting on the surface of the water. During filling and spin-up, the tank was covered with cellophane to minimize wind stresses. The tank rotated counterclockwise and the disk rotated in the same direction. The dense salty water for the plume was dyed green for visualization and injected via a perforated copper pipe ring that encircled the disk. The pipe was covered in foam to evenly distribute the dense fluid. In contrast to the experiments of Bush and Woods, a dense plume was injected at the surface rather than a buoyant plume injected on the bottom. For each experiment, the dense water was injected in 10 - 12 bursts of 20 - 30 seconds each. As discussed in Bush and Woods, there are several "instability times" that define the number of eddies created from the plume. The injection time was limited to the first instability time so that the eddies could be more easily observed. The multiple injections also simulated the more realistic sporadic atmospheric forcing that would create EDW. A conductivity probe measured the water column to get density profiles during the experiments. A camera was mounted on the rotating table that photographed the development of the plume from a sideview. A camera mounted above the table that rotated at the same rate videotaped the plume from above. A syringe pump was used to inject blue dye lines on isopycnals so that upwelling could be observed. Figures 3 and 4 show plan and side views of the experimental setup respectively. Figures 5,6, and 7 show photographs of the experimental setup. The variable parameters in the experiments were the flow rate of the dense water Q $(10 - 37.5\ cc\ s^{-1})$, the reduced gravity of the dense water g'_s $(1.1 - 100\ cm\ s^{-2})$, the disk rotation $\Delta\Omega$ $(0.149 - 3.190\ rad\ s^{-1})$, the buoyancy frequency N $(0.205 - 1.022\ rad\ s^{-1})$, and Coriolis parameter f $(0.8 - 2\ rad\ s^{-1})$.

Figure 4 also shows how the height of the eddies was determined. Geostrophy states that the azimuthal velocity V is

$$V = \frac{g'}{f}\frac{dh}{dx}, \tag{6}$$

where dh/dx is the slope of the isopycnals and is equivalent to S as depicted in Figure 4. Rearranging and scaling g' as N^2H, V as $r\Delta\Omega$ as in Narimousa and Maxworthy, and using Equation 5, h_e can be scaled as

$$h_e = SR = \frac{r\Delta\Omega}{2N}. \tag{7}$$

Note the factor of 1/2. The velocity on the bottom was zero. In this scaling, the average velocity, rather than the surface velocity, was the velocity of interest, so this factor was

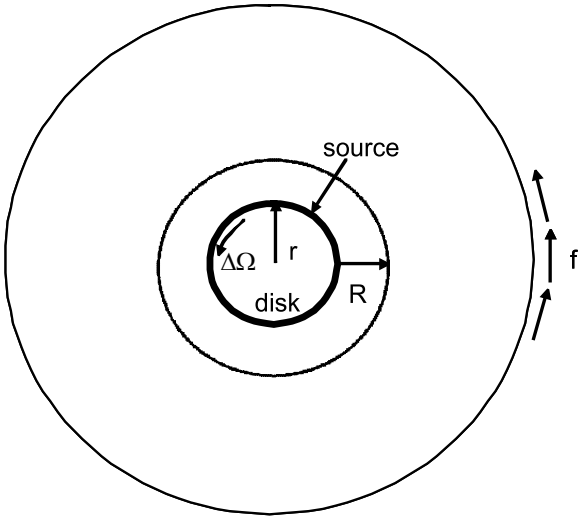


Figure 3: Plan view of the experimental setup.

added. Measurements of the slope from photographs (see Figure 8) show that the above scaling for the slope is accurate within 30%.

3 Theory

The behavior of the dense plume is function of a combination of the vertical buoyancy flux from the plume and the lateral eddy buoyancy flux due the eddies created by the upwelling front. An appropriate scaling needs to be determined for both fluxes. The ratio of these terms should provide insight into the plume behavior. The buoyancy flux is commonly defined as

$$B = g'_s Q, \tag{8}$$

where Q is the volumetric flow rate of the source and g'_s is defined here as

$$g'_s = g\frac{\rho_{source} - \rho_{top}}{\rho_{top}} \tag{9}$$

where ρ_{source} is the density of the source water for the plume and ρ_{top} is the density at the surface of the tank, where the plume is injected.

The lateral eddy buoyancy flux can be written as

$$M = \overline{v'b'} \times 2\pi\frac{2r + R}{2}h_e \tag{10}$$

where $\overline{v'b'}$ is the lateral buoyancy eddy flux per unit area. In this equation, the area through which the flux acts is defined as the height of the eddies created via upwelling (see Figure 4) times the perimeter of the circle through which the flux acts. Here this perimeter is defined as the average between the circle created by the disk and the circle created by the eddies (of radius R - see Figure 3) generated at the edge of the disk. Using a flux-gradient relationship,

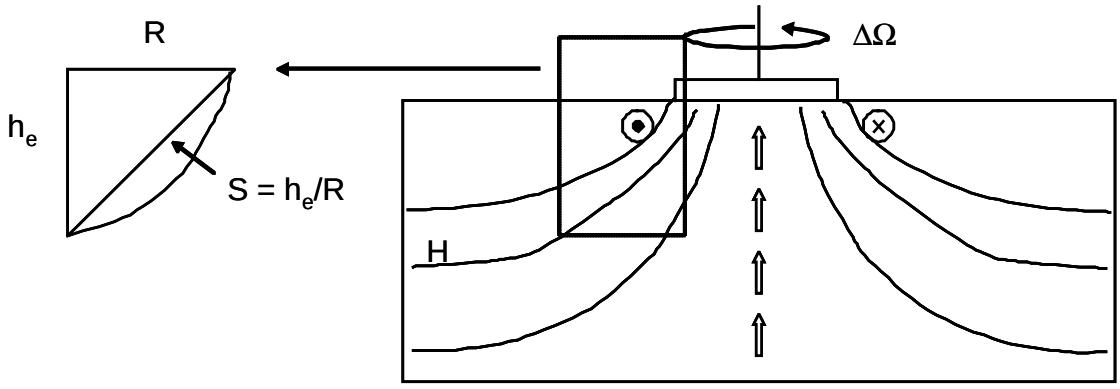


Figure 4: Side view of the experimental setup. The section shown indicates how the height of the eddies was scaled.

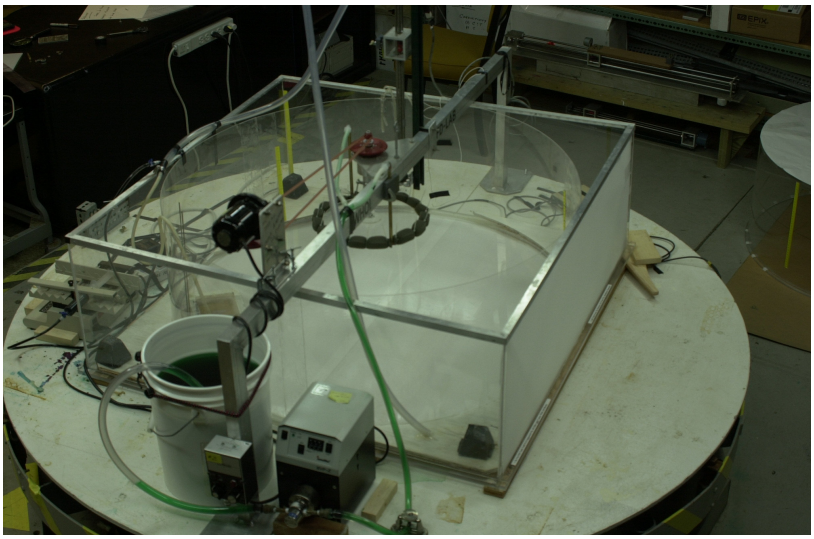


Figure 5: Overhead view of experimental setup. In the foreground the pump that delivers the source can be seen.

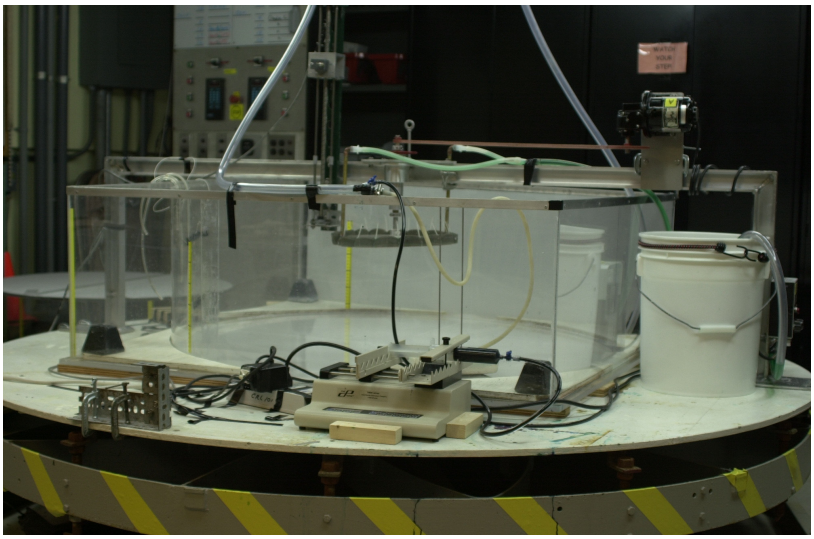


Figure 6: Side view of experimental setup. In the foreground the syringe pump that was used to inject dye lines for visualization of the upwelling can be seen.

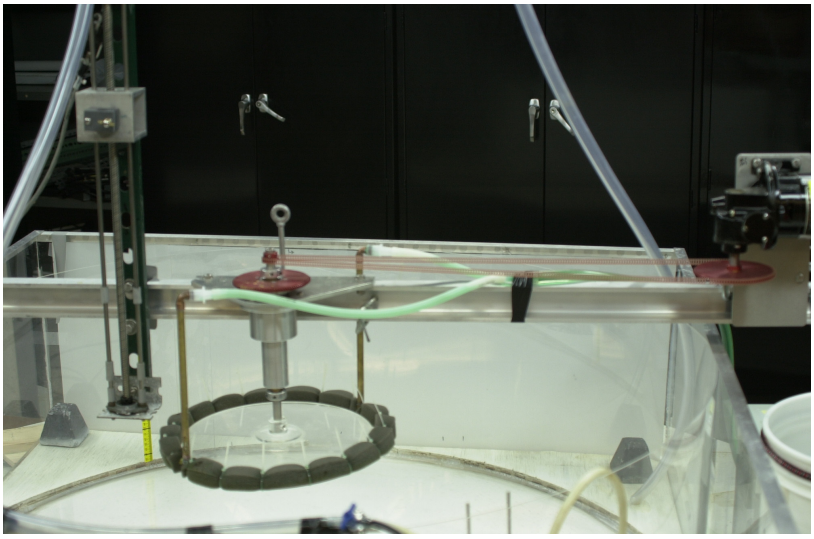


Figure 7: View of the disk that is rotated to create upwelling and the piping that delivers the source. The profiling apparatus for the conductivity sensor is seen to the left of the disk.

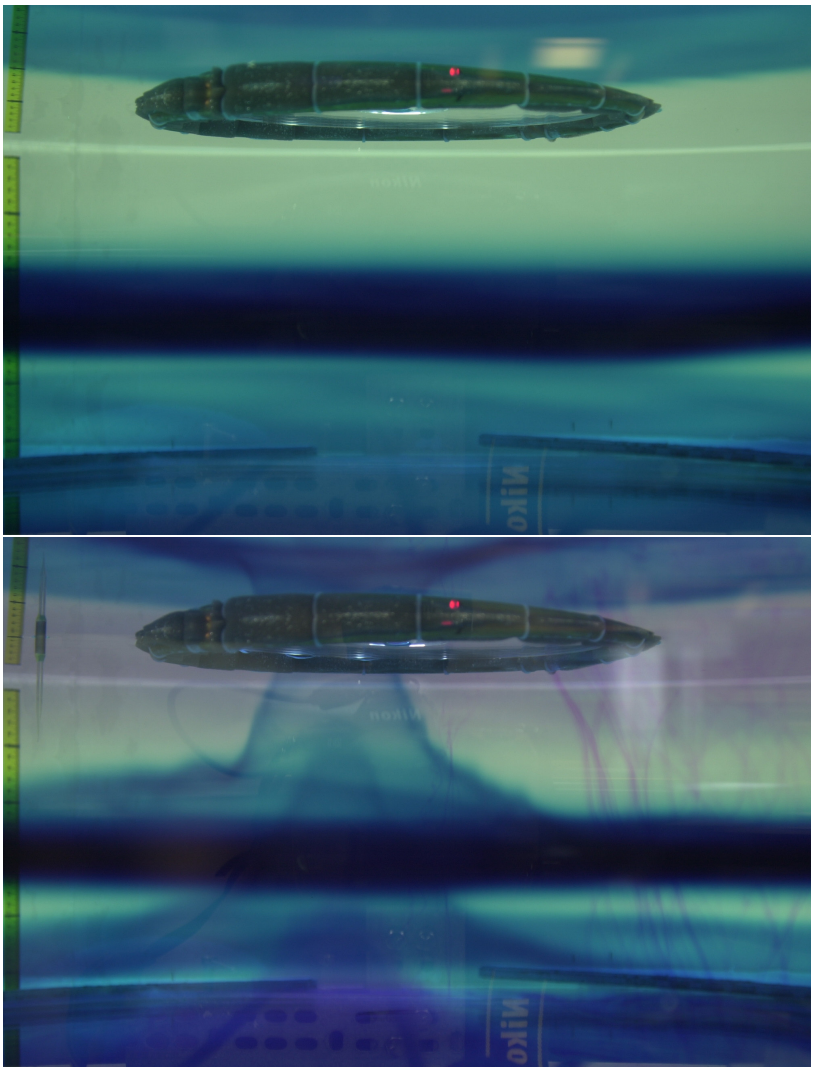


Figure 8: Upwelling in the current experiment. The upper photo shows the isopycnals before the disk starts rotating and the lower panel shows the sloping of the isopycnals.

$$\overline{v'b'} = -K\nabla b, \tag{11}$$

where K is an eddy diffusivity. The eddy diffusivity can be described as

$$K = cVR, \tag{12}$$

where c is an eddy efficiency parameter. Cenedese et al. and others found that the value of the eddy efficiency parameter ranged between 0.01 and 0.06[17]. In the current experiments a value of $c = 0.05$ was used. ∇b can be scaled as $-\Delta b/R$, where b is assumed to change over the lengthscale R. Assuming Δb scales as N^2H (or g') and using Equation 12, the lateral buoyancy flux per unit area is

$$\overline{v'b'} = cVN^2H \tag{13}$$

Using Equations 5 7, and 13 and the scaling for V from Narimousa and Maxworthy (including a factor of 1/2 as stated above), Equation 10 can be rewritten as

$$M = \frac{c\pi}{4}\left(r\Delta\Omega\right)^2 NH \quad 2r + \frac{NH}{f} \quad . \tag{14}$$

In the current experiments the effect of the ratio B/M on the plume behavior was examined:

$$\frac{B}{M} = \frac{g'Q}{\frac{c\pi}{4}\left(r\Delta\Omega\right)^2 NH \quad 2r + \frac{NH}{f}}. \tag{15}$$

It was hypothesized that when $B/M < 1$, there would be a significant lateral buoyancy flux would overwhelm the vertical buoyancy flux; when $B/M = 1$, the source and lateral eddy flux have similar strength so they are both relevant for plume dynamics; and when $B/M > 1$, the plume would behave as in the experiments of Bush and Woods due to the negligible value of the lateral eddy flux. It was also hypothesized that an effective vertical buoyancy flux, which would account for lateral eddy flux could be defined as

$$B* = B - M, \tag{16}$$

which can be rearranged to give

$$B* = B \quad 1 - \frac{M}{B} \quad . \tag{17}$$

4 Results

First, the plume behavior was qualitatively analyzed. Overhead video of the experiments showed a distinct difference in the size of the eddies between the experiments with and without upwelling (with all other parameters held constant). First, the base case with no upwelling is examined. Figure 9 shows the plume (in green) after the first injection of dense fluid (upper photo) and after five injections (lower photo). The eddies formed

from the injection can clearly be seen in the upper photo. After several injections a thin layer formed. Figure 10 compares the base case with no upwelling (upper photo) with the upwelling case (lower photo). The eddies created by the upwelling affect the plume by causing it to mix as it sinks, resulting in a thicker mixed layer whose midpoint is not as deep as that formed with no upwelling. The eddies essentially increase the entrainment of the ambient fluid into the plume as it descends.

The measured density profiles provided a mechanism for a quantitative analysis of the intrusion of the plume. In the current experiments, the depth of the midpoint of the intrusion, the average density of the intrusion, and the thickness of the intrusion were investigated as a function of B/M. Figure 11 shows a typical density profile after the injection of the plume (without upwelling). The plume creates a step in the background density profile, which would be expected based on comparison with the upper photo of Figure 10. Figure 11 also shows the depth that Equation 1 predicts the plume will descend to. Note that this distance was assumed to start at the bottom of the mixed layer that formed on the surface, perhaps due to wind stresses or surface stresses exerted by the disk. The density of the step was used to define a g' for the step (with the respect to the density at the surface). Figure 11 also shows how the thickness of the step, ΔZ, was defined. Figure 12 shows a typical density profile for the same B as in Figure 11, but now with upwelling. The depth predicted by Bush and Woods is still shown. It is apparent that the plume did not descend as deep as in the case without upwelling. The intrusion was also much thicker. Comparison with the lower photo in Figure 10 confirms that the plume mixed more as it descended when upwelling was present. Figure 13 combines the typical profiles of the background, the intrusion without upwelling (plume), the background with upwelling, and the intrusion with upwelling (upwelling). The important thing to note is the differences in the structure of the density step with and without upwelling. Another important fact is that the background density profile does not change dramatically once the upwelling begins.

Figure 14 shows the measured depth of the midpoint of the intrusion plotted in dimensionless form using the scaling of Bush and Woods (Equation 1). Note that for the points where there was no upwelling, M was taken to be 1, although in reality it is 0 and the value of B/M at these points is infinite. The points with the lowest values of B/M were discarded because the midpoint of the intrusion was less than the depth of the mixed layer at the surface after upwelling started but before the plume was injected. Hence it was not possible to discern between the mixed layer and the intrusion. The figure shows the depth of the mixed layer for each experiment. An attempt was made to determine a relationship between the depth of the midpoint and B/M. The relationship needed to asymptote to a constant value as B/M went to $+\infty$ in order to match the results of Bush and Woods and tend to zero as B/M got smaller. The data was fit to a form

$$\frac{Z_n N}{(B/L)^{1/3}} = C_1 exp\left(-C_2 M/B\right), \tag{18}$$

where C_1 and C_2 are constants that were determined to be 2.6 and 0.02, respectively. The value of 2.6 is the asymptote that the data approaches as B/M becomes infinite, which is the same condition as the experiments of Bush and Woods. The value of 2.6 is smaller than their average value of 3.0, but within their stated range. The current value is also very

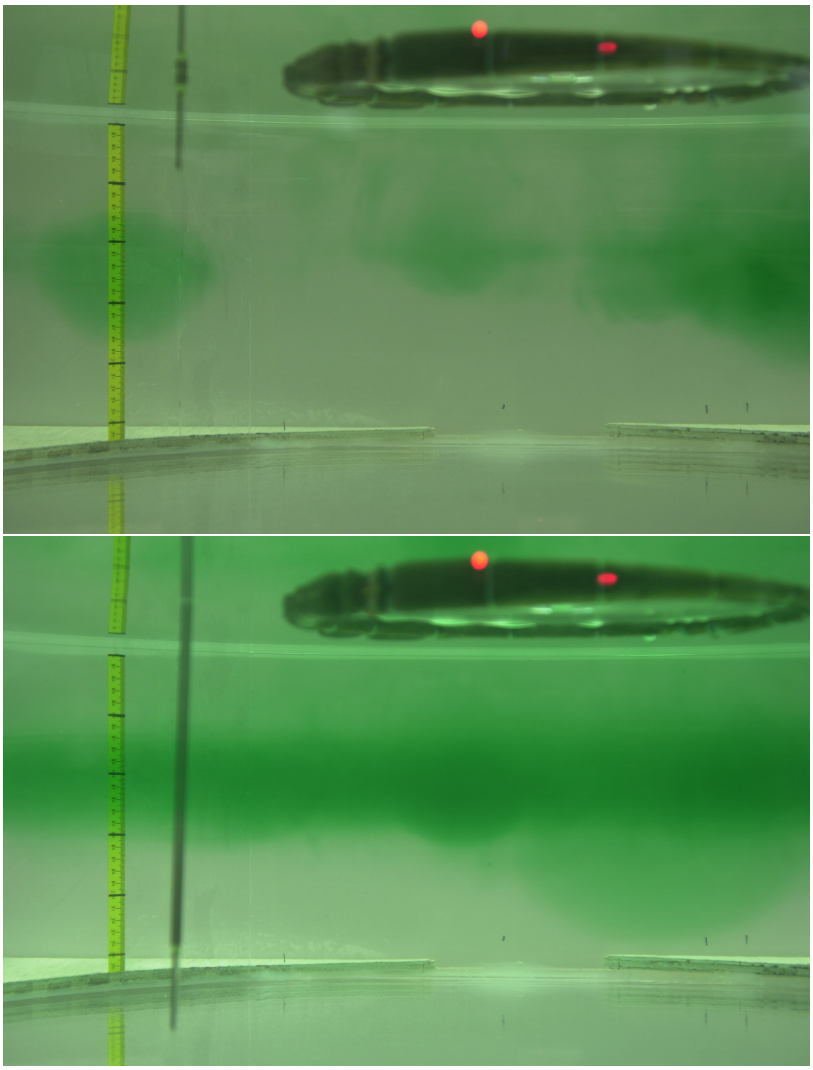


Figure 9: Behavior of a plume with no upwelling ($B = 400\ cm^4 s^{-3}$) after one injection (upper photo) and five injections (lower photo).

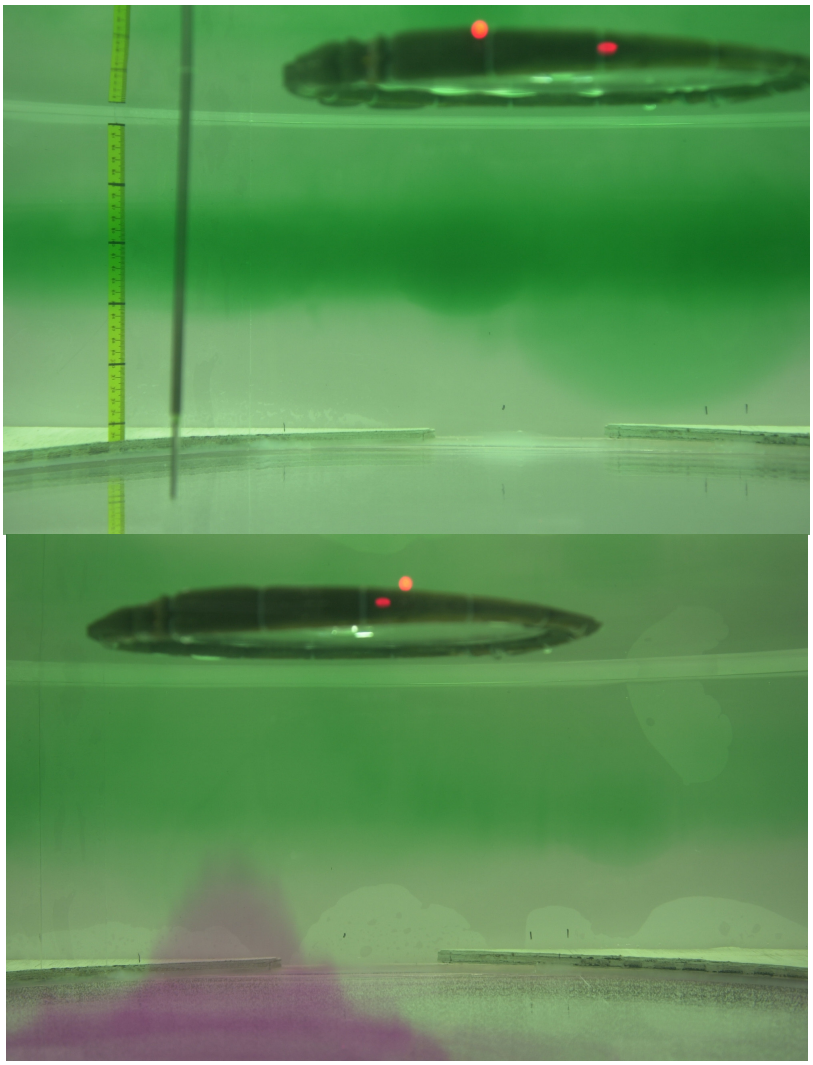


Figure 10: Comparison of plume behavior in the plume only case (upper photo) and the plume in the presence of upwelling (lower photo). All the parameters were the same between the two photos and they were both taken after five injections. $B/M = 0.02$ in the lower photo. The purple potassium permanganate on the bottom of the lower photo shows the upwelling.

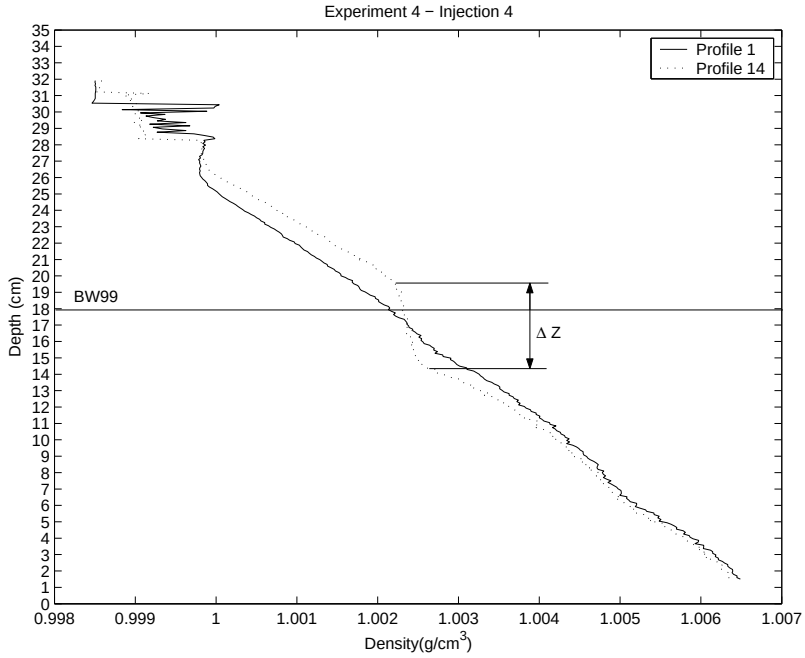


Figure 11: Density profile for a plume with no upwelling ($B = 220\ cm^4 s^{-3}$). The definition of the density step is shown as well as the depth of the intrusion as predicted by Bush and Woods. Profile 1 is the background density profile before injection. In this experiment the mixed layer is seen to extend to 26 cm.

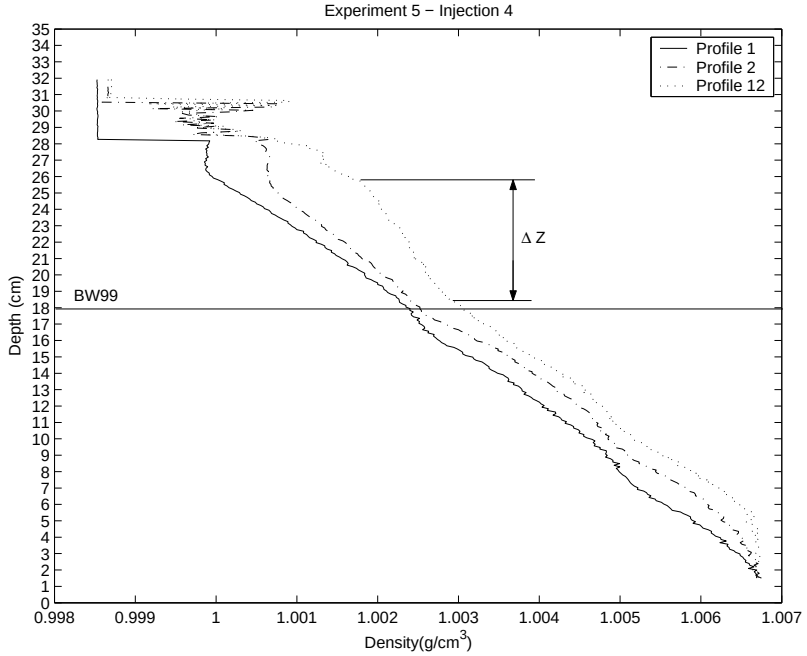


Figure 12: Density profile for a plume with upwelling ($B/M = 0.02$). Profile 1 is the background density profile before injection and prior to starting the disk. Profile 2 is the density profile after the disk was started but prior to injection.

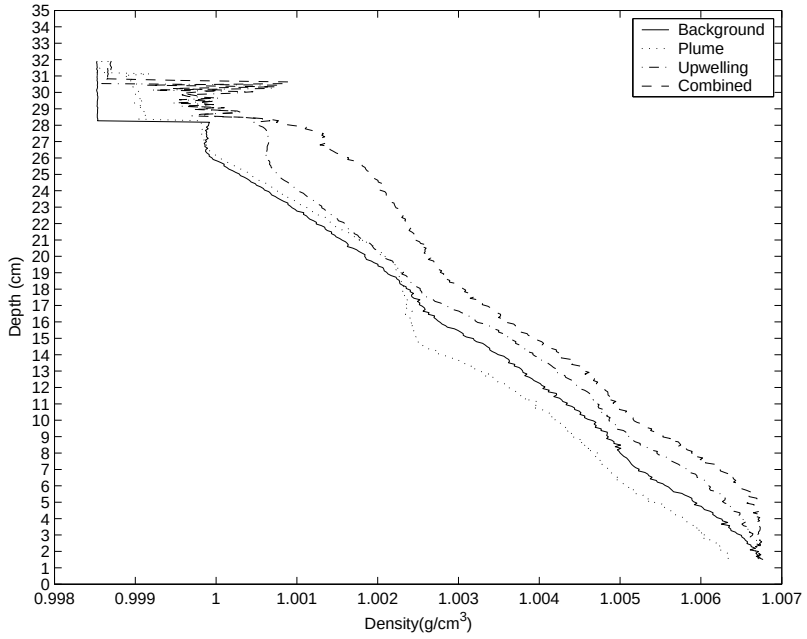


Figure 13: Summary of typical density profiles seen in the current experiments.

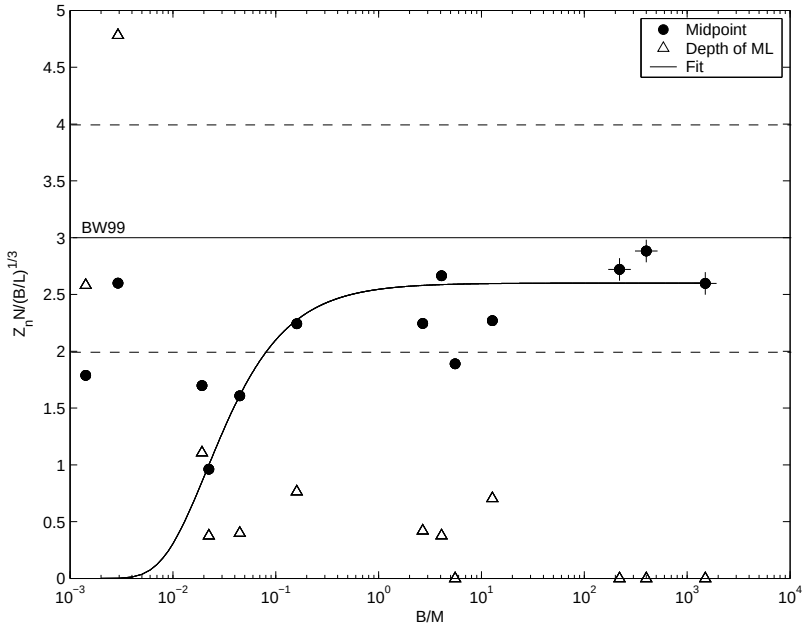


Figure 14: Depth of the intrusion due to the plume. The circles indicate the measured depth and the triangles indicate the depth of the mixed layer. The predicted value of the dimensionless depth from Bush and Woods is shown as a solid line with its margin of error showed as dashed lines. The data points with crosses indicate the points with no upwelling, and hence an infinite B/M.

similar to the value of 2.5 found by Wright and Wallace for the nonrotating case. In general the depth of the intrusion decreased with decreasing B/M, which is compatible with the hypothesis that as M increases the plume mixes more quickly and thus does not descend as deeply.

Figure 15 shows g_i' of the step normalized by the scaling of Bush and Woods (Equation 3). While there was much scatter in this data, there was a general increase in the dimensionless reduced gravity with B/M. This was the expected trend because as B/M decreased, the plume did not descend as deep, thus the density at its neutral level was closer to the density at the surface.

Figure 16 shows the thickness of the step normalized by the predicted height of the eddies created by the upwelling from Equation 7. Note that the experiments that had no upwelling are excluded from this figure because h_e is undefined in these experiments. For $B/M < 1$, the normalized thickness was essentially constant. This is consistent with the eddies from the upwelling having a great effect on the plume behavior in this regime. As B/M increased, the eddies had less of an effect on the structure of the intrusion. This is confirmed by examining the dimensionless thickness as a function of B and M individually. Figure 17a shows that there was no obvious dependence on B. Figure 17b shows that the trend observed in Figure 16 is due to changes in M. Hence, the thickness of the step was correlated with the strength of the lateral eddy buoyancy flux for $B/M < 1$.

Lastly, the proposed effective buoyancy flux from Equation 17 was evaluated. The

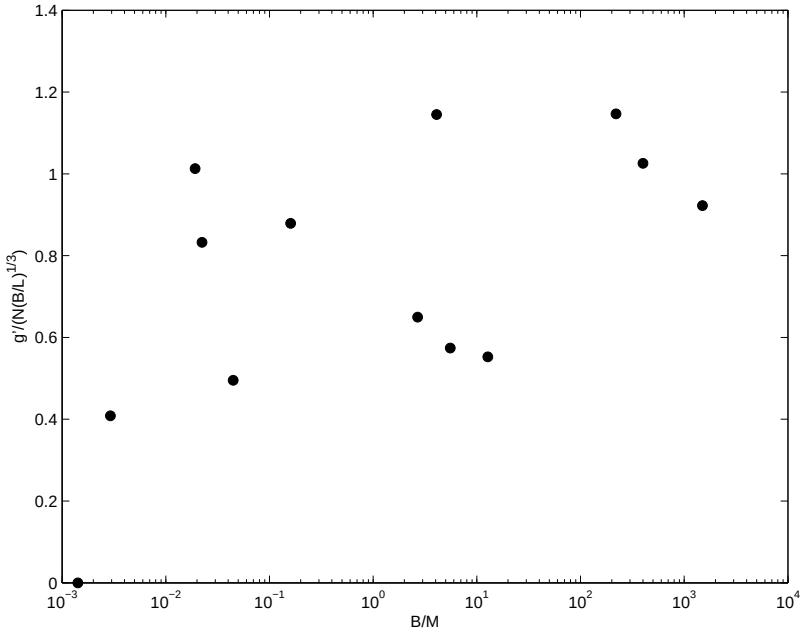


Figure 15: Reduced gravity of the intrusion with respect to the surface density.

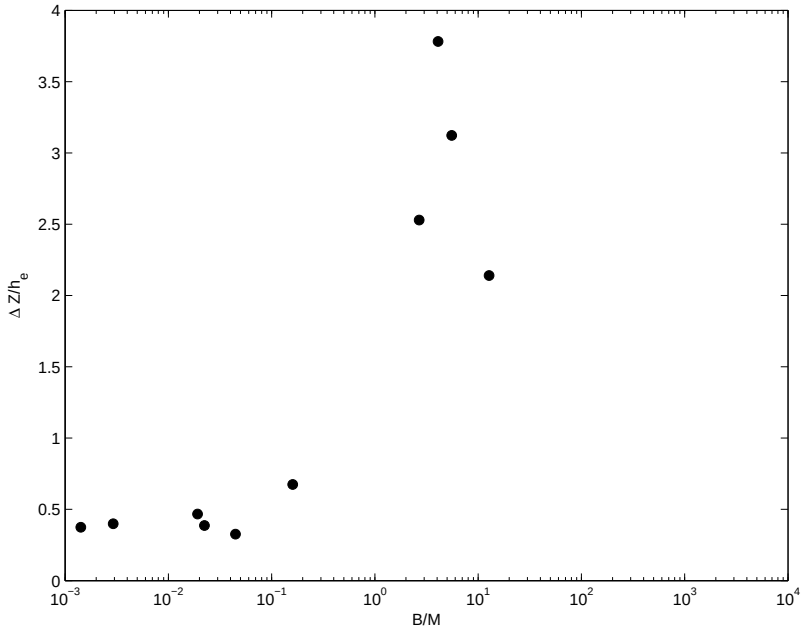


Figure 16: Thickness of the intrusion.

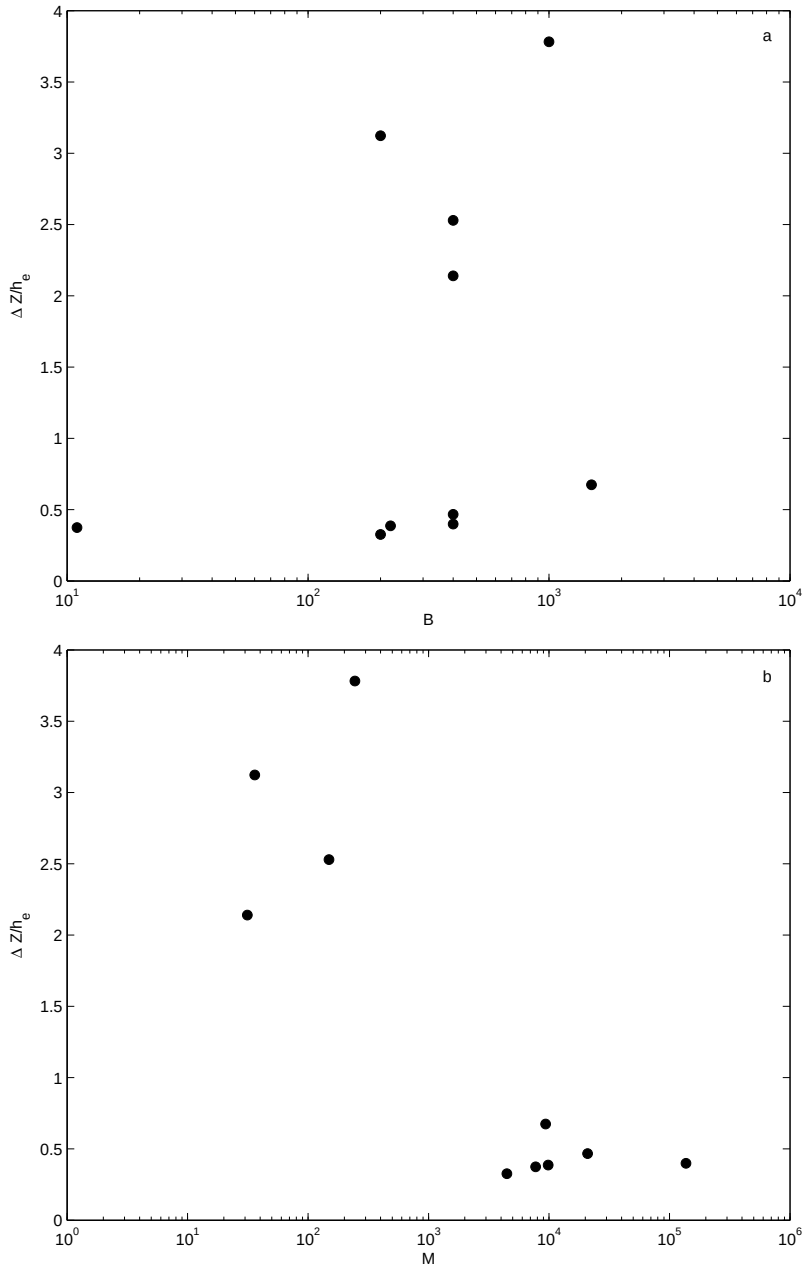


Figure 17: Thickness of the intrusion as a function of B and M independently.

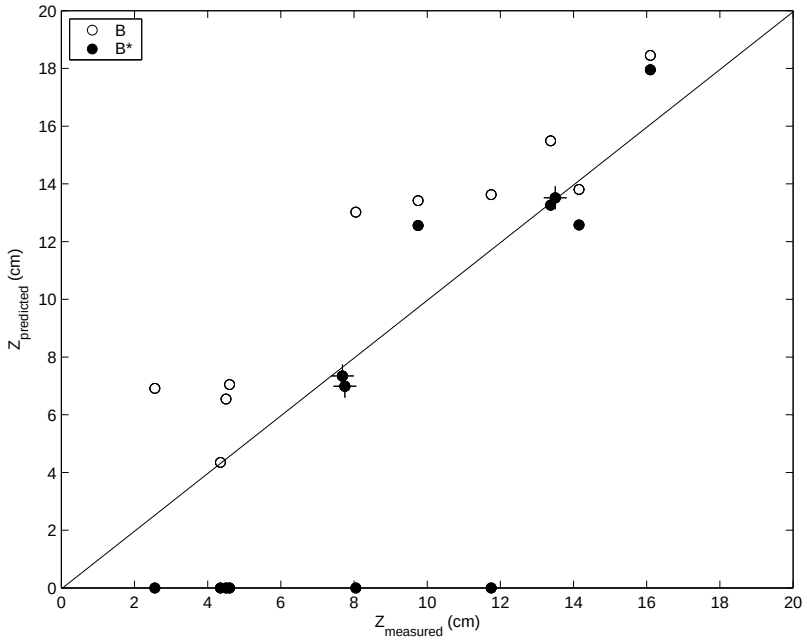


Figure 18: Comparison of the predictions of the intrusion depth using the standard buoyancy flux as defined in Equation 8 and the effective buoyancy flux defined in Equation 17. The dashed line indicates the line of equality between the measured and predicted values. The data points with crosses indicate the points with no upwelling and hence $B* = B$. Note that the points along the x-axis are due where the effective buoyancy flux is not valid.

theoretical intrusion depth was computed using the scaling of Bush and Woods with a coefficient of 2.6 instead of 3.0 to better fit the current data (see Equation 1). Figure 18 compares how the scaling performs with B as defined in Equation 8 and $B*$ as defined in Equation 17. Note that one limitation of this formulation is that it is not valid for $B/M < 1$, which limits its application when the lateral eddy buoyancy flux is expected to dominate. For the remaining points, in general there is an improvement when using the prediction with the effective buoyancy formulation.

5 Discussion

To gain insight into the formation of EDW, the ratio B/M in this region was computed. From the heat flux in Figure 1, the buoyancy flux per unit area was determined to be 10^{-7} $m^2 s^{-3}$. From this figure, the surface area (SA) over which this flux acts was measured to be 1000 km by 200 km. A typical velocity for the Gulf Stream is 1 ms^{-1} and a typical depth over which the gulf stream acts is 1 km. The Rossby radius of deformation in this region is on the order of 100 km. The value of N/f in this region ranges between 30 and 100, so typical stratifications are between $N = 3 \times 10{-}3$ and 10^{-2} s^{-1}. The cross sectional area (CSA) through which the lateral eddy buoyancy flux moves is the length of this zone times the depth of the gulf stream. So

$$B = B/area \times SA = 2 \times 10^4 m^4 s^{-3}$$
$$M = cVN^2H \times CSA = 2 - 5 \times 10^6 m^4 s^{-3}$$

giving a value of $B/M = 4 \times 10^{-3}$ to 10^{-2}. Based on the current experiments, the EDW zone falls into the regime where eddies from the Gulf Stream dominate and the properties of the plume were not important. The effective buoyancy flux formulated above could not be used here because of the low value of B/M. While the EDW falls into this regime, interestingly even the Bush and Woods scaling, which does not account for the lateral eddy fluxes, underestimated the depth of the the EDW, predicting a depth of $80 - 270\ m$ (depending on N/f). The EDW is typically seen between 200 and 500 m, so towards the lower end of the ratio of N/f, predictions approach the observations. Note that in the current experiments, N/f ranged between 0.1 and 1. Another possible explanation for the shallower intrusion depth predicted could be due to the use of a seasonally averaged buoyancy flux on the surface. Using a time averaged buoyancy flux through an entire experiment reduced the B/M values from the lab by a factor of 100, putting the estimated values from the ocean in the middle of the range investigated in the experiments. It is not clear if this is more physically appropriate, as the instantaneous plume properties may be what controls the intrusion behavior, not the time averaged plume properties.

6 Conclusion

Laboratory experiments were conducted to investigate behavior of a dense plume in a rotating linearly stratified fluid in the presence of eddies created by a baroclinic unstable front. The purpose of the experiments was to simulate the formation of Eighteen Degree Water to gain insight into the properties of this mode water. The vertical and lateral eddy buoyancy fluxes were estimated and their ratio, B/M, used to determine properties of the intrusion formed from the plume. The depth of the intrusion was found to be smaller in the presence of eddies from the baroclinic front. An effective buoyancy flux was defined that improved predictions of the depth of the intrusion as long as $B/M > 1$. Further analysis of the scaling for B and M is necessary before application to Eighteen Degree Water can be successfully made.

7 Acknowledgements

I would like to thank the GFD program for giving me the opportunity to be here this summer, Claudia Cenedese for the idea behind the experiments and all her time and guidance, John Marshall for help with the scaling and the theory, and Keith Bradley for setting up the rotating table, fixing all the things that kept breaking while I was running these experiments, and conversation during those long hours in the lab. I would also like to thank the 2-m rotating table for only trying to kill me once. Lastly I want to give mad props to my officemates and the other fellows for constant entertainment and helping me get through the requisite all-nighters to complete this project, when Diet Dr. Pepper alone would not suffice. Marshall gets a special shout out for being my lab buddy and for being such a help with my figures.

References

[1] K. Speer and E. Tziperman, "Rates of water mass formation in the North Atlantic ocean," J. Phys. Ocean. **22**, 93 (1992).

[2] Y.-O. Kwon and S. C. Riser, "Observations of the north atlantic subtropical mode water," First ARGO Science Workshop (2003).

[3] T. Joyce, A. Hernandez-Guerra, and W. Smethie, "Zonal circulation in the NW Atlantic and Caribbean from a meridional World Ocean Circulation Experiment hydrographic section at 66 degrees W," J. Geophys. Res. **106**, 22095 (2001).

[4] H. Rouse, C. Yih, and H. Humphreys, "Gravitational convection from a boundary source," Tellus **4**, 201 (1952).

[5] B. Morton, G. Taylor, and J. Turner, "Turbulent gravitational convection from maintained and instantaneous sources," Proc. Roy. Soc. London **234**, 1 (1956).

[6] S. Wright and R. Wallace, "Two-dimensional buoyant jets in stratified fluid," J. Hydraul. Div. Am. Soc. Civ. Eng **105**, 1393 (1979).

[7] Y. Noh, H. Fernando, and C. Ching, "Flows induced by the impingement of a 2-dimensional thermal on a density interface," J. Phys. Ocean. **22**, 1207 (1992).

[8] C. Ching, H. Fernando, and Y. Noh, "Interaction of a negatively buoyant line plume with a density interface," Dyn. Atmos. Oceans **19**, 367 (1993).

[9] H. Fernando and C. Ching, "Effect of background rotation on turbulent line plumes," J. Phys. Ocean. **23**, 2125 (1993).

[10] J. Lavelle and D. Smith, "Effects of rotation of convective plumes from line segments sources," J. Phys. Ocean. **26**, 863 (1996).

[11] H. Fernando, R. r. Chen, and B. Ayotte, "Development of a point plume in the presence of background rotation," Phys. Fluids **10**, 2369 (1998).

[12] J. Bush and A. Woods, "Vortex generation by line plumes in a rotating stratified fluid," J. Fluid Mech. **388**, 289 (1999).

[13] K. Helfrich and T. Battisti, "Experiments on baroclinic vortex shedding from hydrothermal plumes," J. Geophys. Res. **96**, 12511 (1991).

[14] K. Speer and J. Marshall, "The growth of convective plumes at sea-floor hot-springs," J. Geophys. Res. **96**, 12511 (1991).

[15] S. Narimousa and T. Maxworthy, "2-layer model of shear-driven coastal upwelling in the presence of bottom topography," J. Fluid Mech. **159**, 503 (1985).

[16] S. Narimousa and T. Maxworthy, "Coastal upwelling on a sloping bottom - the formation of plumes, jets, and pinched-off cyclones," J. Fluid Mech. **176**, 169 (1987).

[17] C. Cenedese, J. Marshall, and J. Whitehead, "A laboratory model of thermocline depth and exchange fluxes across circumploar fronts," J. Phys. Ocean. **34**, 656 (2004).

Experimental study of double diffusive gravity currents under rotation

Marshall Ward

1 Introduction

It is well known that when light and heavy fluid are adjacent to each other, there is an unbalanced pressure gradient that triggers an adjustment process towards equilibrium, so that the light fluid finds its way atop the heavy fluid. This dam-break or lock-exchange problem has been analyzed in great detail for laminar flow (e.g. [1]), though the turbulent situation is perhaps not nearly as well understood. But in either case it is clear that, during adjustment, the light fluid will be pushed upward, with the heavy fluid similarly pushed downward, in the form of a hydraulically-forced gravity current. The process, however, becomes more elaborate when the lateral density gradients vary over long lengthscales. Nearly all large-scale terrestrial phenomena, from river outflow to large-scale oceanic gyres, feel the influence of the Earth's rotation, and this drastically alters the adjustment process. To complicate matters even further, the presence of both heat and salinity distributions in the oceans can lead to buoyantly stable density profiles that are nonetheless double diffusively unstable, and is most prevalent when foreign water intrudes as a gravity current into another part of the ocean. Two potential examples of this are the Mediterranean outflow into the North Atlantic and the various boundary currents of the Arctic Ocean, both of which are intrusions of warm salty water into a cold, fresher background; such intrusions also tend to be steered by the topography for some distance. To estimate the role of these intrusions accurately, it is necessary to understand the underlying processes and how they can interact to influence gravity current formation and propagation.

One of the earliest complete analyses of nonrotating gravity currents is due to Benjamin [2]. A major result of his study is that a frictionless gravity current, once established, will propagate at a constant speed of $c = \sqrt{2g'h_b}$, where g' is the reduced gravity and h_b is the lower fluid depth. It was also shown that, far upstream of the front, the thickness of the lower fluid will occupy only half of the total depth. If friction is introduced, the bottom height may become smaller. Experimental work on nonrotating gravity currents by Rottman and Simpson [3] has shown that, in the early stages, the displacement of the front is linear in time, so that the speed is indeed constant. At later times, speeds begin to slow and the displacement curve obeys a power law of either $t^{2/3}$ if there is an interaction with the end of the tank, or $t^{1/5}$ if viscous effects dominate; the time and nature of such a transition depends on the Reynolds number of the current, though in all cases the viscous profile eventually dominates.

The formation and evolution of a gravity current under the effects of rotation has been studied extensively, and has been reviewed by Griffiths [4]. The classic geostrophic adjustment problem of Rossby [5] was extended to a rotating channel flow by Gill [6], who considered the adjustment of two semi-infinite channels filled with fluid of infinitesimally different heights and separated by a barrier. Using linearized dynamics, it was shown that, for a channel of width larger than, but still comparable to, the radius of deformation, the adjustment process is initially driven by the propagation of gravity waves, just as for a nonrotating fluid. But after an inertial period of $2\pi f^{-1}$, where f is the Coriolis parameter, a semigeostrophic balance is established and the adjustment process is then dominated by the propagation of coastal Kelvin waves away from the location of barrier release. Since these waves travel along the boundaries in opposite directions, with the coast to the right of the direction of propagation (for $f > 0$), they give rise to gravity currents that hug the coast in a similar manner. This analysis has been extended to the case of finite difference of fluid height in a number of later studies (e.g., [7, 8]). The result of Benjamin mentioned earlier was generalized to rotating currents by Stern, Whitehead, and Hua [9], who found that the speeds of these currents are also constant, but that, because certain quantities remain indeterminate without a more rigorous analysis, the estimate for c becomes an upper bound.

Although the dynamics of gravity currents, both in the absence and presence of rotation, are understood in some detail, there has been less focus on how they are influenced by double diffusion. Laboratory experiments have shown that, for fluids with compensating temperature and salinity gradients, double diffusive effects can induce intrusions into each fluid, in the form of interleaving patterns [10, 11]. The effect of shear on the magnitude of the interfacial density flux has also been considered [12]. But direct studies of double diffusive gravity currents are few. Maxworthy [13] provided the first thorough experimental study of such currents, focusing on two-layer experiments of both constant flux and constant volume. He considered currents whose layer density differences were on the order of 10^{-3} to 10^{-2} g cm^{-3}, many times larger than in the study presented here. At early times, these currents exhibited behavior common to strongly inertial nondiffusive currents, such as the formation of a bulbous head and a constant frontal propagation rate (i.e. head speed). The propagation rates also did not differ noticeably from nondiffusive currents of the same density difference. At later times however, the heads would narrow to sharp peaks and, in the case of a finger favorable current, strong convection would begin, carrying heavy fluid downward and leading to the formation of a second, bottom gravity current. The speeds would also drop dramatically at this stage, and in some cases would stop completely.

Yoshida, Nagashima, and Ma [14] approached this problem from a different perspective by noting that a double diffusive mass exchange across the interface should increase the layer density difference over time, and thereby accelerate the gravity currents in some way. They suggest that the layer density differences considered by Maxworthy were too large to notice such a change, and so they looked at the evolution of double diffusive gravity currents with density differences that were at least an order of magnitude smaller. They were able to demonstrate that the double diffusive currents are notably faster than the nondiffusive ones. But because of the indirect method used to specify the density of each layer, they were unable to make a direct comparison between, say, finger favorable and nondiffusive, but otherwise identical, currents. They also showed that the finger favorable currents produced

a secondary bottom current, similar to those observed by Maxworthy.

In this study we follow an approach similar to that of Yoshida et. al. by using density differences on the order of 10^{-4} g cm^{-3}, though we also reduce the fluid height. While this makes our currents more susceptible to frictional effects, the small initial density differences and low speeds also make them highly sensitive to double diffusive fluxes. In Section 2 we describe our laboratory setup and the experiments performed. Sections 3 and 4 contain our results for the propagation rate of the fronts and their lateral spread. Analysis of the results, with a focus on the small-scale structure, is provided in Section 5. Conclusions and suggested areas for future work are discussed in Section 6.

2 Experimental Setup

2.1 Apparatus

Two sets of lock-exchange gravity current experiments were performed at the Woods Hole Oceanographic Institution. Both used an acrylic tank of dimensions 60 cm long, 10 cm wide, and 20 cm high, with a barrier located at the center (30 cm) point. The first set of experiments were done for a nonrotating system, and the setup is denoted schematically on Figure 1. The second set utilized a rotating table, shown in Figure 2.

For each lock-exchange experiment, two solutions of different densities and concentrations were prepared, using salt and sugar as our fast and slow diffusers[1]. The densities for each solution were specified by first adding the amount of solute prescribed by Ruddick and Shirtcliffe [15]. We then refined the density by dilution or addition of solute, based on measurements from an Anton Paar DMA58 density meter. This allowed us to attain the desired density within 10^{-5} g cm^{-3}. All solutions were prepared from storage vats of distilled water, which were given sufficient time to adjust to room temperature.

Each run required 500 ml of a designated "light" fluid of lower density and an equal amount of "heavy" fluid of higher density. Unless specified otherwise, the light fluid was dyed with blue food coloring, whose effect on density was taken into account. The fluids were poured into opposite ends of the tank, which were separated by a movable barrier. We then verified that the fluid heights on the two sides were equal, correcting each level if necessary. For the rotating cases, we allowed for 10 minutes of spinup time to let the system reach equilibrium. After the system settled, the barrier was then quickly but carefully removed, with as little disturbance to the fluid as possible, and the gravity current was allowed to develop.

2.2 Visualization

Two methods of visualization were used during these experiments. For observation of the vertical structure, a minimal amount of blue food color dye (less than 0.1 ml) was added to the lighter fluid to distinguish the individual fluid layers. A shadowgraph, illuminated by a projector light source, allowed for resolution of the small-scale double diffusive features. For the lateral structure of the rotating currents, 2 ml of food color dye was used to highlight the structure of the current profile. In both cases, images were collected with either a

[1]The diffusivities of salt and sugar in water are 1.5×10^{-5} and 0.5×10^{-5} cm^2 sec^{-1}, respectively.

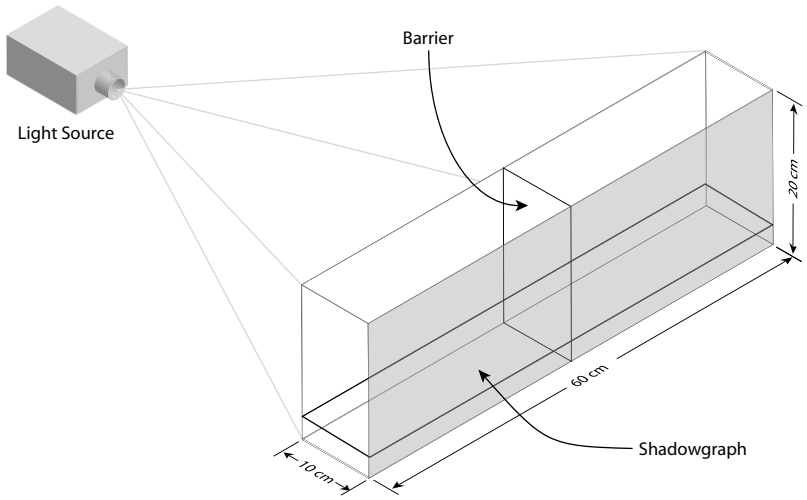


Figure 1: Experimental Setup for Nonrotating Experiments

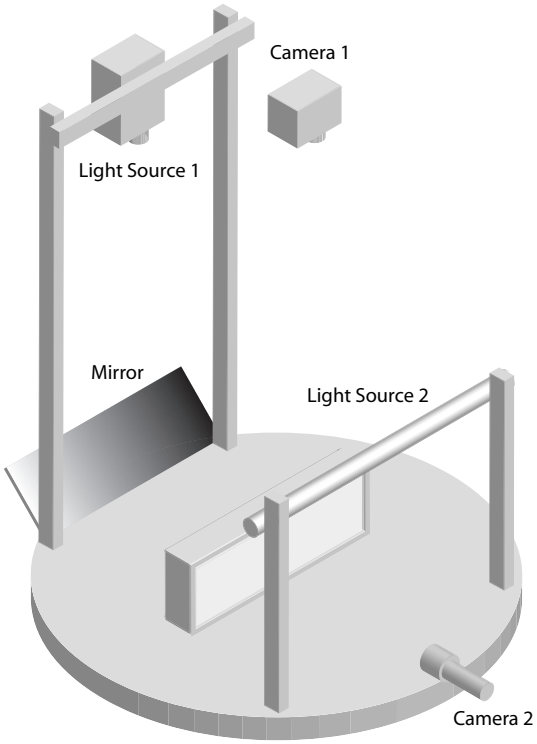


Figure 2: Experimental Setup for Rotating Experiments

black-and-white video camera for lower resolution video capture, or a digital still camera for higher resolution image analysis.

For most experiments, the overhead cameras track the propagation of the gravity current along the upper fluid surface. However, as mentioned before, currents with finger favorable profiles can produce secondary bottom currents that can, in more extreme cases, become the dominant gravity current and may tend to contain most of the dyed fluid. Our visualization methods will therefore track the bottom current in these situations. For such cases, it was confirmed that the top and bottom currents propagate at the same speeds however, so that for the purposes of speed comparison it is sufficient to track either current.

2.3 Experimental Parameters

Table 1 shows the physical parameters of the experiments to be discussed. The first row refers to the double diffusive process that one would expect from each experiment. For example, if we consider an experiment where light sugar water will flow above heavy salt water, then we would expect fingers to form. Such currents are referred to as finger favorable (hereafter FF) currents. Similarly, a flow with light salt water over heavy sugar water would be expected to form a diffusive layer, and the currents are said to be diffusive layer favorable (hereafter DF). We would not expect any significant diffusive behavior if both fluids contain only salt, or only sugar.

Table 1: Experimental Parameters

Parameter	Values
Double Diffusive Tendency	Fingers, Diffusive Layers, or Single Component
Rotation Rate (f)	0.0 or 1.0 rad $\sec^{-1}$
Prospective Difference ($\Delta\rho$)	5.0×10^{-4} or 1.2×10^{-4} g cm^{-3}
Mean Density (ρ)	1.02 or 1.04 g cm^{-3}
Fluid surface height (h)	1.7 cm

Table 2: Physical and dimensionless parameters ($U \approx 0.03$–$0.15\,\mathrm{cm\ sec}^{-1}$)

Parameter	$\Delta\rho = 0.0005$	$\Delta\rho = 0.00012$
$c = \sqrt{g'h}$	$0.91\,\mathrm{cm\ sec}^{-1}$	$0.44\,\mathrm{cm\ sec}^{-1}$
$L = c/f$	0.91 cm	0.44 cm
$\mathrm{Fr} = U/c$	0.03–0.16	0.07–0.34
$\mathrm{Ro} = U/fL$	0.03–0.16	0.07–0.34
$\mathrm{Re} = Uh/\nu$	5.1–25.5	5.1–25.5

We also consider the effects of rotation, density difference between layers, and solute concentration of the fluids. We specify the rotation rate in terms of the Coriolis parameter, $f = 2\Omega$, where Ω is the true rotation rate of the table. The density difference between layers is established during preparation of the fluids, in the manner described before. The mean density is equivalent to solute concentration. The fluid height h results from the use

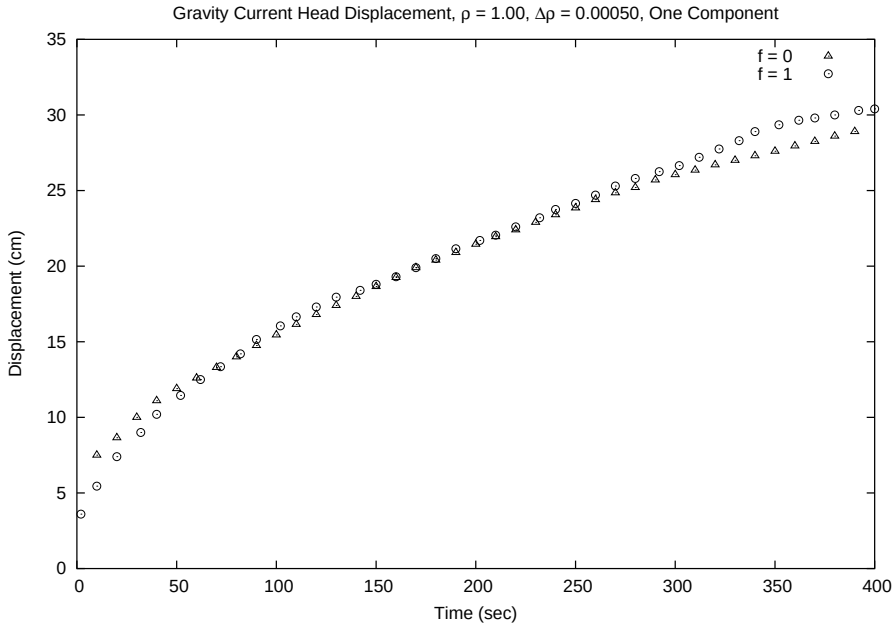


Figure 3: Effect of rotation on a nondiffusive current with a large $\Delta\rho$

of 1000 ml of fluid in our particular tank. A complete set of the experiments performed is shown in Tables 3 and 4, shown at the end of this manuscript.

Table 2 gives some estimates for the dynamical and dimensionless parameters of our experiments, based on empirical observations of gravity current head speeds of 0.03–0.15 cm sec^{-1}. The dynamical parameters, c and L, are the estimates for the (shallow water) gravity wave speed and Rossby radius of deformation, respectively. The dimensionless parameters are estimates for the Froude, Rossby, and vertical Reynolds numbers of our experiments. We see that Froude and Rossby numbers are small, and so we should expect some sort of geostrophically or quasigeostrophically balanced flow, with rapid adjustment by the internal gravity waves. Reynolds numbers are, however, only modestly large, so we can only presume that viscous effects will play a significant role in the dynamics of the gravity currents.

3 Frontal Propagation

In this section we provide a quantitative description of the propagation rate of the interfacial front of a two layer double diffusive gravity current, and its variation in response to the following effects: rotation, double diffusive tendency, density difference between layers, and solute concentration.

3.1 Rotation

We first consider the response of a current to rotation. The head speed of a single component current with and without rotation is shown in Figure 3. The plot shows that there is close agreement between the two curves during the intermediate period of 50–200 sec, but there

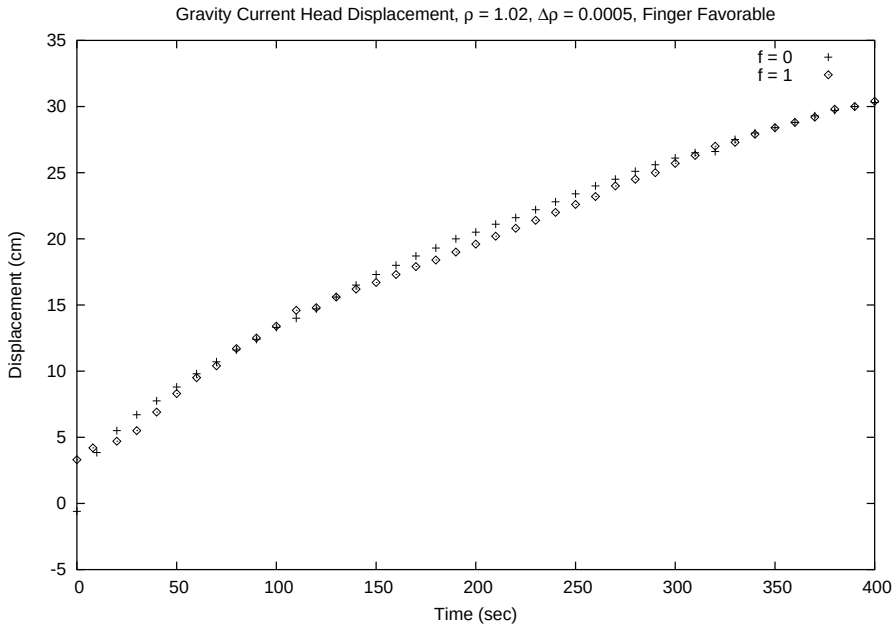


Figure 4: Same as Figure 3, for a FF current

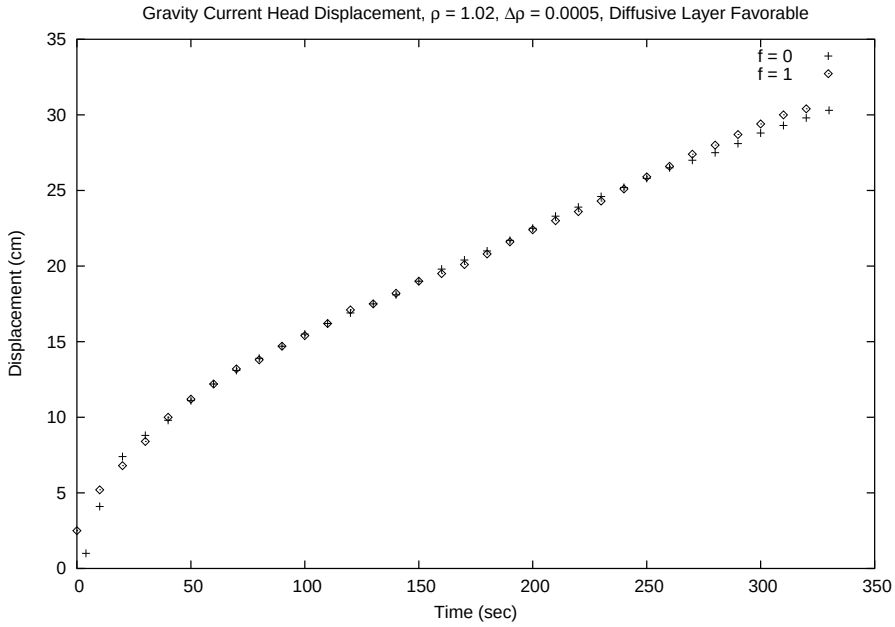


Figure 5: Same as Figure 3, for a DF current

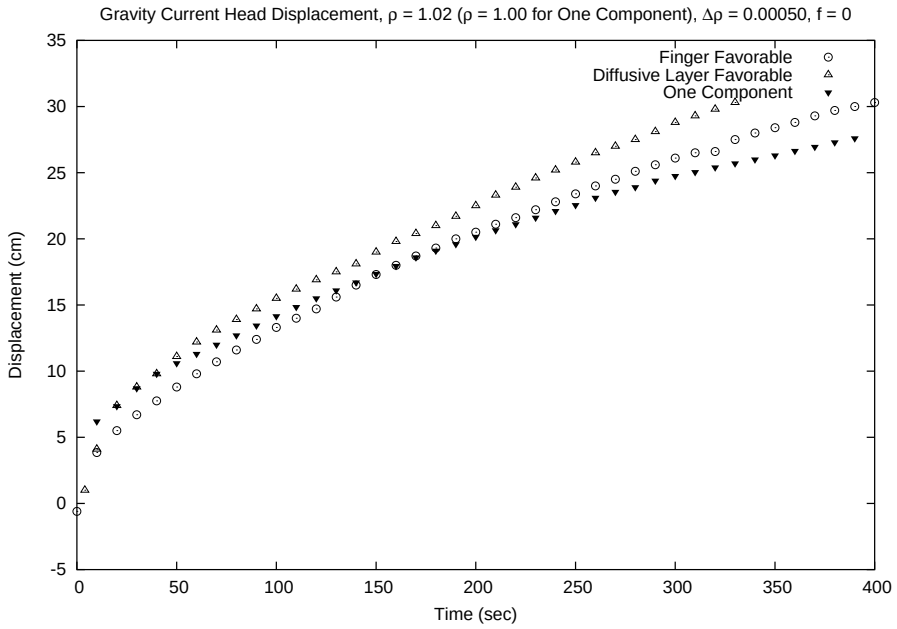


Figure 6: Effect of double diffusion on a nonrotating gravity current with large $\Delta\rho$

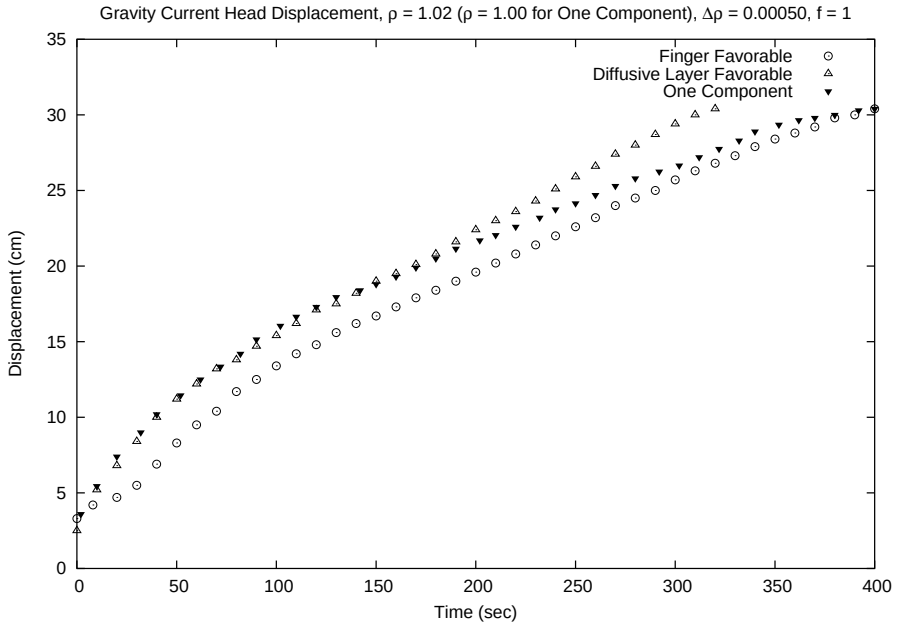


Figure 7: Same as Figure 6, with rotation

is a noticable discrepancy at later times, particularly as the gravity current reaches the end of the tank.

Figure 4 shows that there is also close agreement between the rotating and nonrotating curves for the FF gravity currents. The agreement is even better between the DF currents, shown in Figure 5. In both cases, the curves appear to be piecewise linear, with transitions occurring after about one minute. This is somewhat expected, since frictionless currents have been shown to propagate at constant speeds [2, 9]. But it is surprising that this particular state would dominate for so long, since our Reynolds numbers are not very large and viscosity would be expected to have a substantial influence on the propagation, as shown by Rottman and Simpson [3].

In all three scenarios, we find that there is reasonable agreement between the head speeds of rotating and nonrotating gravity currents. This is an indication that the fluid velocities perpendicular to the current motion, and resulting Coriolis forces along the current, are small and have little effect on frontal propagation. Consequently, the dynamics along the current are largely inertial, with the acceleration driven by a lateral pressure gradient and modified by viscous forces. And since speeds are much larger in the direction of the current, the Coriolis forces are significant in the direction perpendicular to the current propagation and we would expect a crossflow geostrophic balance. This is demonstrated by the formation of a boundary current in all rotating cases, such as those in Figures 13–18.

3.2 Double Diffusive Tendency

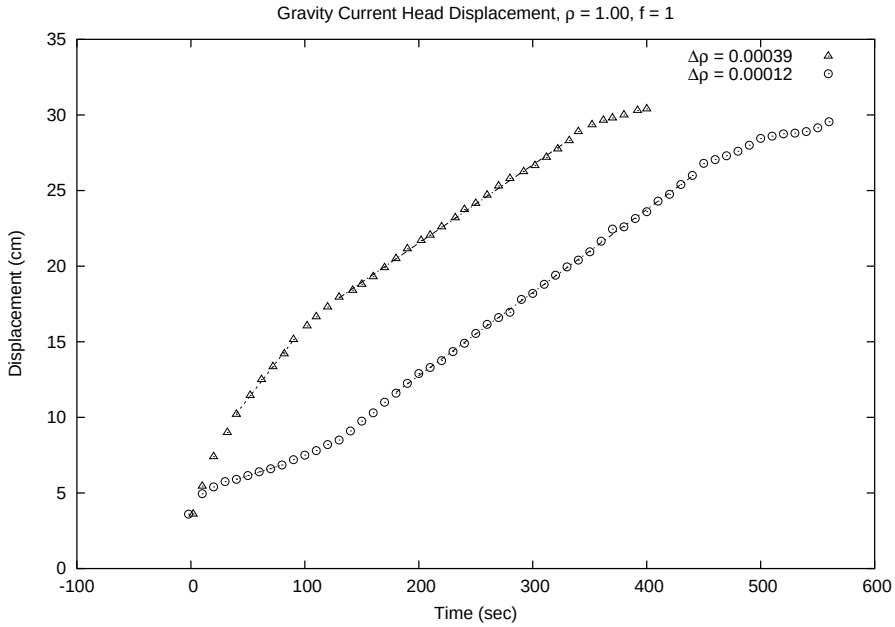
Figures 6 and 7 demonstrate, in the absence and presence of rotation, respectively, the behavior of gravity currents that have different double diffusive profiles but are otherwise equivalent. In both plots, the double diffusive gravity currents are significantly faster than the nondiffusive ones. For the rotating cases, there are intermediate regimes where the speed profile appears to be approximately linear.

If we focus on Figure 7 but appeal to the linear fits of the intermediate regimes of the same curves that appear in Figures 8, 9, and 10 (namely, the curves for large $\Delta\rho$ in each plot), then we see that the DF currents are 33% faster than the nondiffusive currents, while the FF currents are only about 13% faster. The greater speeds of the double diffusive currents are most likely a consequence of the increasing density difference between layers over time. This can in principle lead to a stronger acceleration and larger head speed.

3.3 Density Difference Between Layers

The most dramatic behavior in head speed is exhibited by variation of the initial density difference between layers. Hereafter, we will restrict our attention to data from rotating experiments. We also use curve fitting to estimate the head speeds for some of the observed linear regimes. While this method is somewhat subjective because it is not always clear where the linear regimes begin and end, it provides a quantitative estimate for comparative purposes.

Figures 8, 9, and 10 demonstrate two general responses to changes of initial layer density difference, the first being that the strong initial surge associated with large density differences is much weaker, if it occurs at all, when the density difference is reduced. The



$\Delta\rho$ (g cm^{-3})	U (cm sec^{-1})	Δt (sec)
0.00039	0.097	40–90
0.00039	0.052	130–330
0.00012	0.024	40–80
0.00012	0.055	180–440

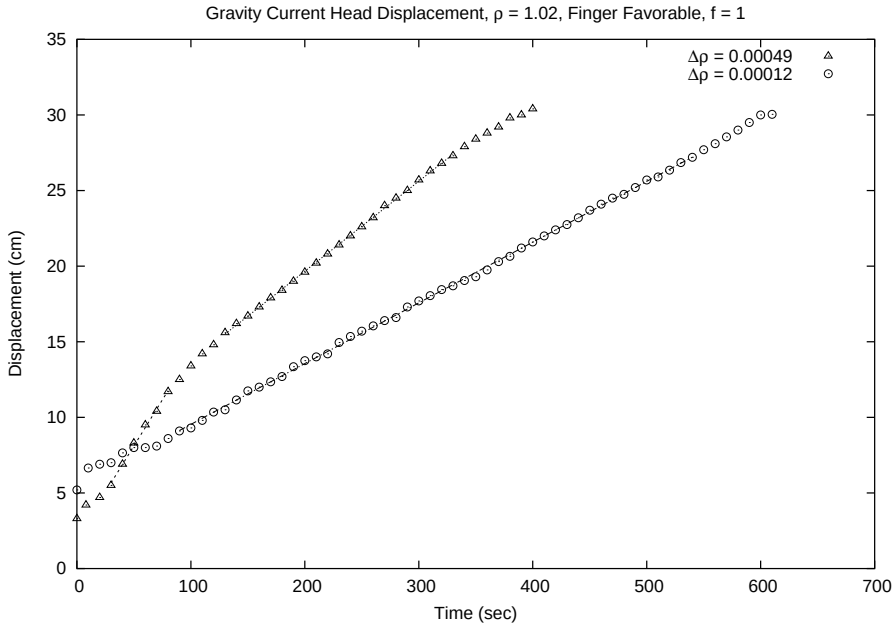
Figure 8: Effect of layer density difference on a nondiffusive current with rotation

second, more predictable, result is that head speeds become larger as density differences are increased.

Figure 8 illustrates a dramatic feature which is observed only in nondiffusive currents. We see that when the density difference is very small, the head speed is initially much less than the gravity current with a larger density difference. But after about two minutes, this slower current begins to separate from the wall and the speed increases dramatically. Evidence of this separation can be seen in Figure 16. The table corresponding to Figure 8 shows that the speeds of the currents in the two cases are in fact comparable. Some speedup after separation is expected, since there should be less viscous drag along the sidewall. But it is surprising that the speedup would be so dramatic, and that it would match the speed of the current with a larger density difference.

Such separation is not observed among the double diffusive currents. Figure 9 shows the evolution of FF gravity currents with large and small initial density differences. At early times (until about the first minute), the current with a large density difference is relatively fast, while the current of small difference appears to be much slower. At later times we have more predictable results, with the larger $\Delta\rho$ producing currents that are about 50% faster than the smaller $\Delta\rho$, as we see from the table below Figure 9.

Figure 10 and the corresponding table show that the DF currents exhibit some behavior



$\Delta\rho$ (g cm^{-3})	U (cm sec^{-1})	Δt (sec)
0.00049	0.122	30–80
0.00049	0.059	130–330
0.00012	0.040	90–540

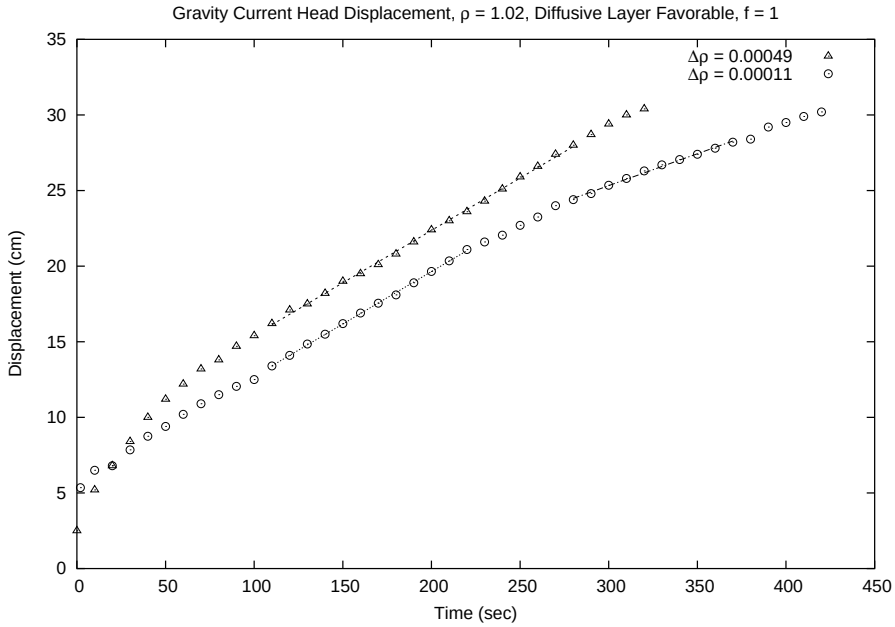
Figure 9: Effect of layer density difference on a FF current with rotation

that is similar to the FF case. However, there are some notable distinctions between the FF and DF currents when the density difference is small. After the first two minutes, this current's speed is nearly the same as the one with a large initial density difference. It is only after about five minutes that there is an abrupt reduction in speed, so that the current with a large difference is nearly 66% faster than the current with a small difference.

In both cases the currents have greater head speeds when the initial density differences are larger, but the speeds do not scale like $\sqrt{\Delta\rho}$, which one might expect from dimensional arguments or comparison to theoretical results [2, 9]. One possible reason for this is that the head speed is nonzero even when $\Delta\rho$ is initially zero, which has been demonstrated in previous studies [16], and is a consequence of increasing density difference over time.

3.4 Solute Concentration

Figure 11 and the corresponding table show that, after a two minute adjustment period, the FF current with a higher solute concentration is more than 50% faster than the current of lower concentration. Both currents also appear to undergo a transition at roughly the same time. The effect is similar for the DF current, shown in Figure 12, which appears to approach its dominant speed more rapidly, occurring after about a minute. The higher concentration current is also nearly 50% faster than the lower concentration one, and resembles



$\Delta\rho$ (g cm^{-3})	U (cm sec^{-1})	Δt (sec)
0.00049	0.069	110–280
0.00011	0.069	110–220
0.00011	0.042	280–370

Figure 10: Effect of layer density difference on a DF current with rotation

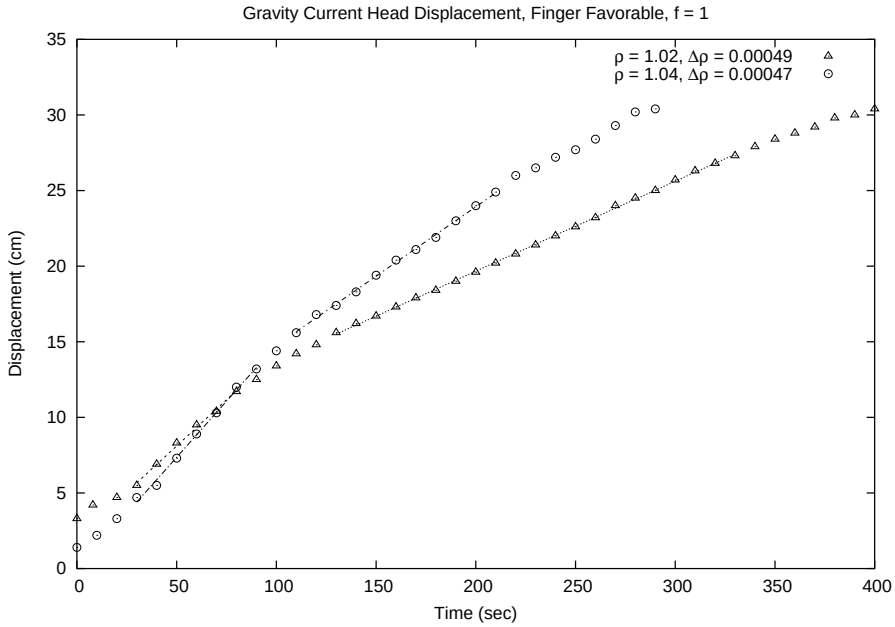
the behavior of the FF current.

Our results demonstrate that solute concentration can be as significant of a factor as the density difference when the density differences are very small. Just as a stronger concentration gradient will typically drive a greater diffusive flux, in a Fickian sense, it is not surprising that double diffusive fluxes across the interface will also be intensified in the presence of stronger solute gradients.

4 Current Structure

We now focus on the lateral structure of the currents. In all cases where rotation is involved, coastal currents are formed, shown in Figures 13, 14, and 15 for a density difference of approximately 5×10^{-4} g cm^{-3}. The most distinguishing feature of the double diffusive currents (shown in Figures 14 and 15) is that the shear-like protrusions associated with the nondiffusive fronts are no longer present. Such protrusions were first studied by Stern [17], and a more recent investigation has been conducted by Stern and Chassignet [18]. While the diffusive currents often exhibit a wavy interface, there is no sign of the sharp crests observed in the nondiffusive current. There is also a noticable lateral spread, resulting in a broadening of the double diffusive currents as compared to the nondiffusive ones.

This spread is even more pronounced for currents where the density difference has been



ρ (g cm^{-3})	U (cm sec^{-1})	Δt (sec)
1.02	0.122	30–80
1.02	0.059	130–330
1.04	0.148	30–90
1.04	0.092	110–210

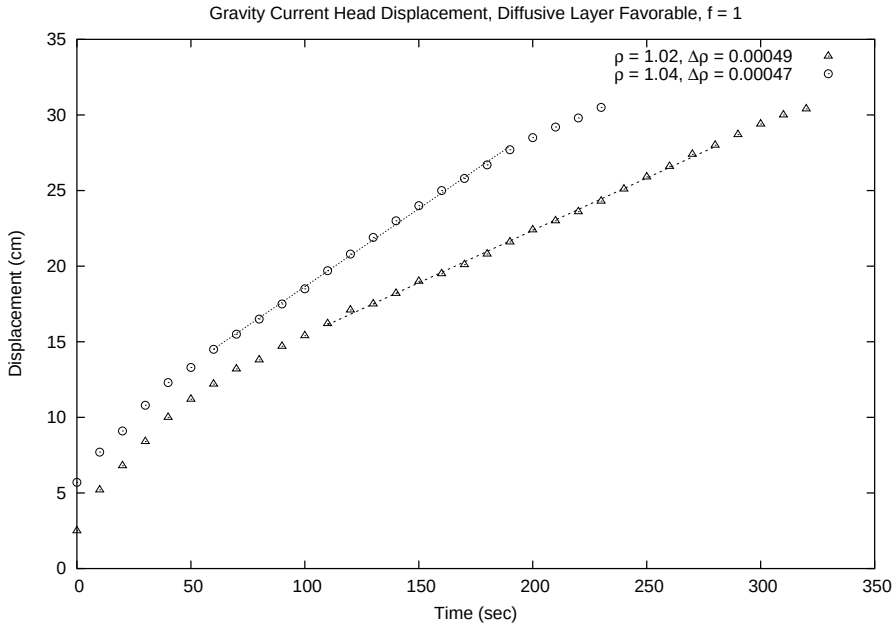
Figure 11: Effect of solute concentration on head speed of a FF current

reduced by a factor of four, shown in Figures 16, 17, and 18. It also appears that these currents tend to retain any vorticity that forms during barrier release for a much longer time; this is most evident in Figure 17. And as discussed earlier, the nondiffusive currents with small density difference tend to separate from the wall, observable in Figure 16. This is not seen in the double diffusive currents, and may be a consequence of the lateral spread. In any case, it is clear that double diffusive fluxes, and the changing density difference with time, are responsible for these differences.

5 Internal Flow Structure

Without an accurate means of measuring the density field over time, it is difficult to diagnose the influence of double diffusion on the gravity currents with great confidence. Nonetheless, observations provide clues about the flow structure and the dynamical processes behind them.

We first consider the flow of a DF current. From shadowgraph observations, we observe that, for a gravity current propagating to the right, there is a strong clockwise shear in the vicinity of the diffusive layer, with weaker counterclockwise shears within each layer. A vertical cross section showing the flow profile is given in Figure 19, and photographs



ρ (g cm^{-3})	U (cm sec^{-1})	Δt (sec)
1.02	0.069	110–280
1.04	0.103	60–190

Figure 12: Effect of solute concentration on head speed of a DF current

of a typical current are shown in Figure 22. We expect that diffusive transport across the layer will carry salt downward, making the fluid just above the layer lighter and the fluid just below it heavier. If the rate of diffusive transport is reasonably slow, then we would not expect an onset of convective mixing but instead an organized, buoyantly induced upward transport of lighter fluid along the interfacial slope in the top layer, and a downward transport of heavier fluid in the bottom layer. The increased flow rate would then lead to a greater gravity current speed. In general though, this induced flow will probably be greater than the increased frontal propagation rate, so that the current will reverse direction near the front. This would result in the counterclockwise shear flows within each layer, though the lower layer shear would be significantly reduced by bottom friction. Such shears would also be expected to induce mixing within each layer.

The evolution of a gravity current under a FF stratification is markedly different. The initial protrusion of the gravity current is illustrated in Figure 20. There should be an initial period of finger formation along the front, with freshened salt water moving upward and a salt-sugar mix moving downward. After these fingers penetrate deeply into each layer, they will eventually become unstable and induce local convective mixing near the top and bottom of the fluid. This will eventually lead to the creation of a very fresh solution atop, with the original sugar and salt solutions inbetween, and a very heavy salt-sugar solution at the bottom, shown schematically in Figure 21. The actual flow will not have the clear boundaries shown in this figure; the density variation will be more continuous

Figure 13: View from above of a rotating, nondiffusive gravity current, four minutes after release of the barrier; each tick on the scale represents one centimeter

Figure 14: Same as Figure 13, for a FF current

Figure 15: Same as Figure 13, for a DF current

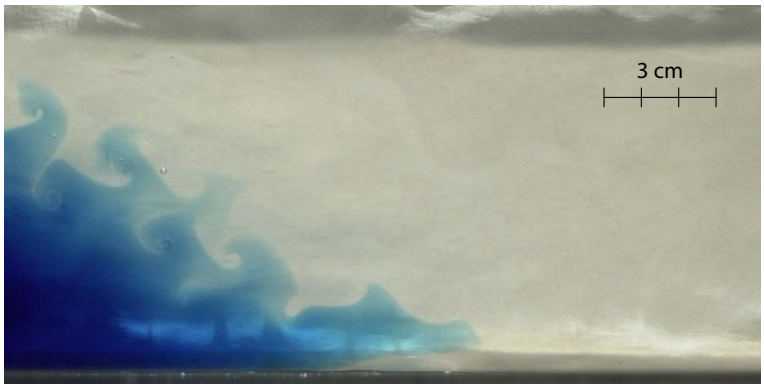


Figure 16: View from above of a rotating, nondiffusive gravity current, four and a half minutes after release of the barrier.

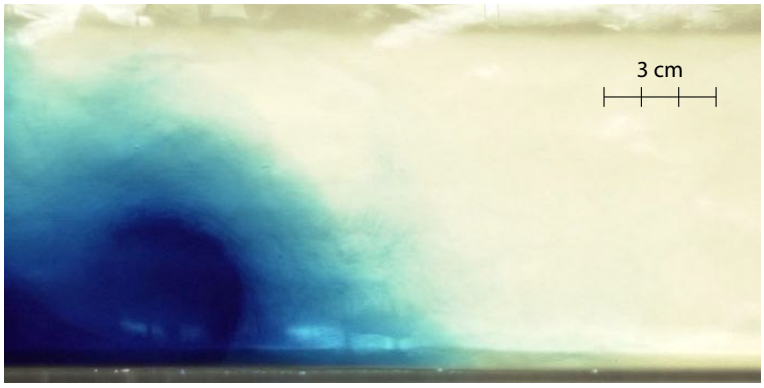


Figure 17: Same as Figure 16, for a FF current

Figure 18: Same as Figure 16, for a DF current

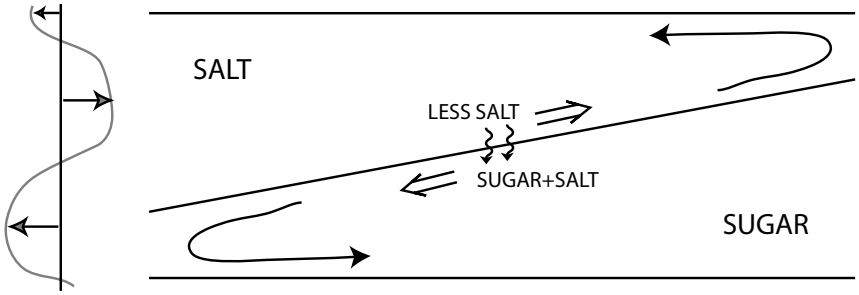


Figure 19: Flow structure of a DF current

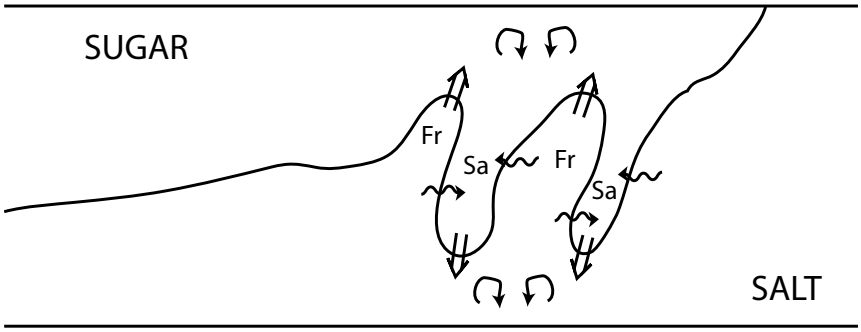


Figure 20: Early intrusion of a FF current

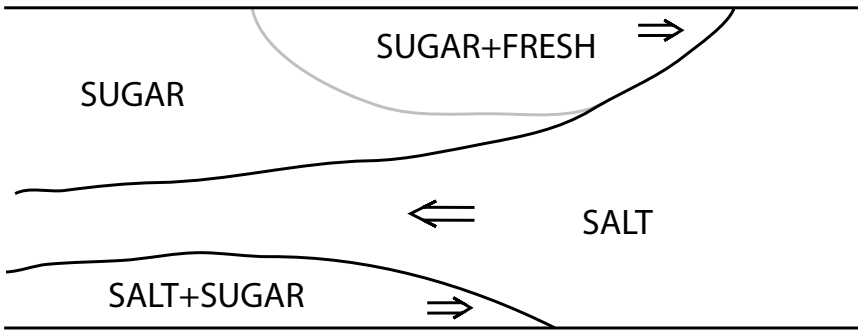


Figure 21: Schematic of a FF current after induced mixing

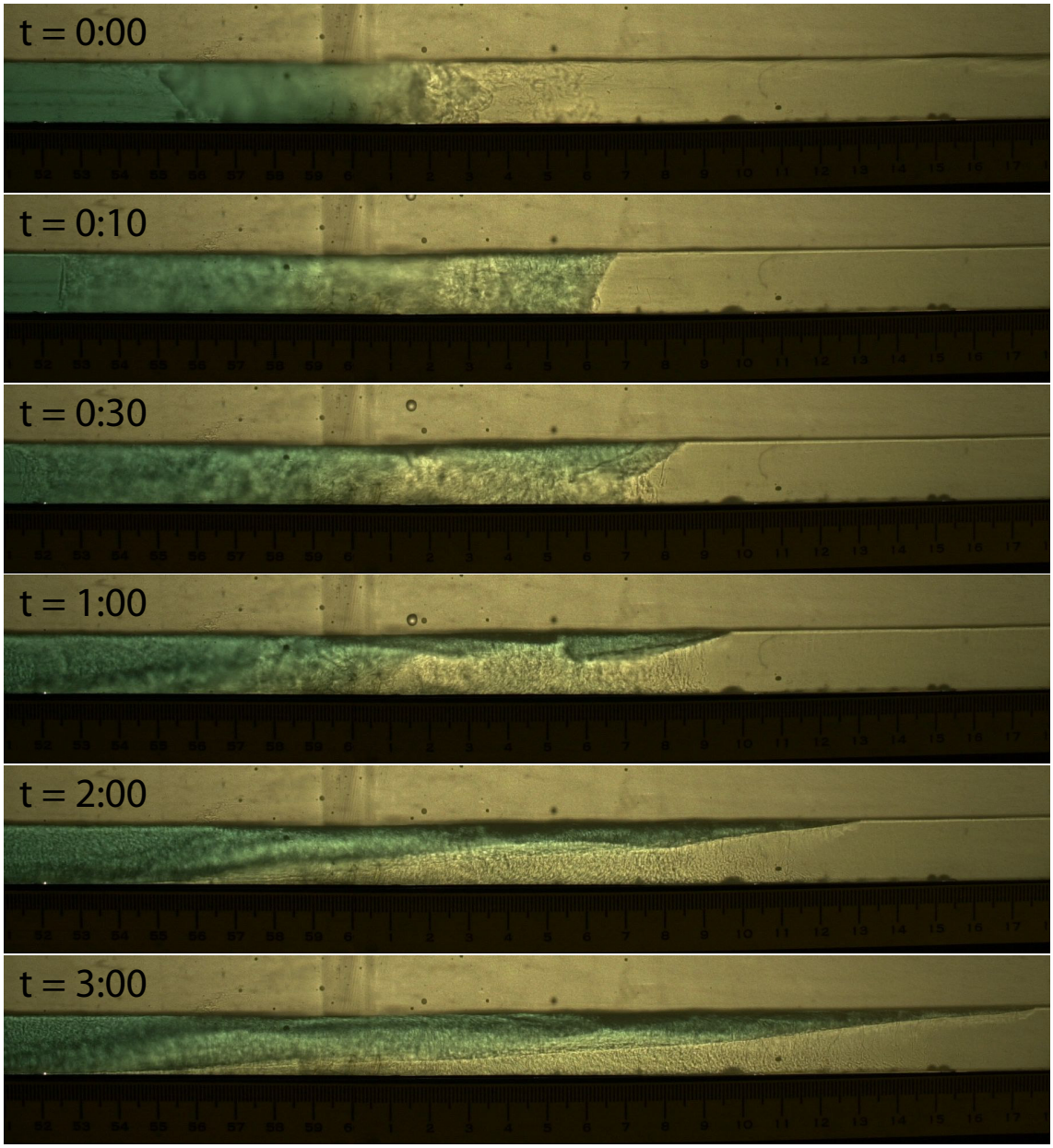


Figure 22: Evolution of a DF gravity current

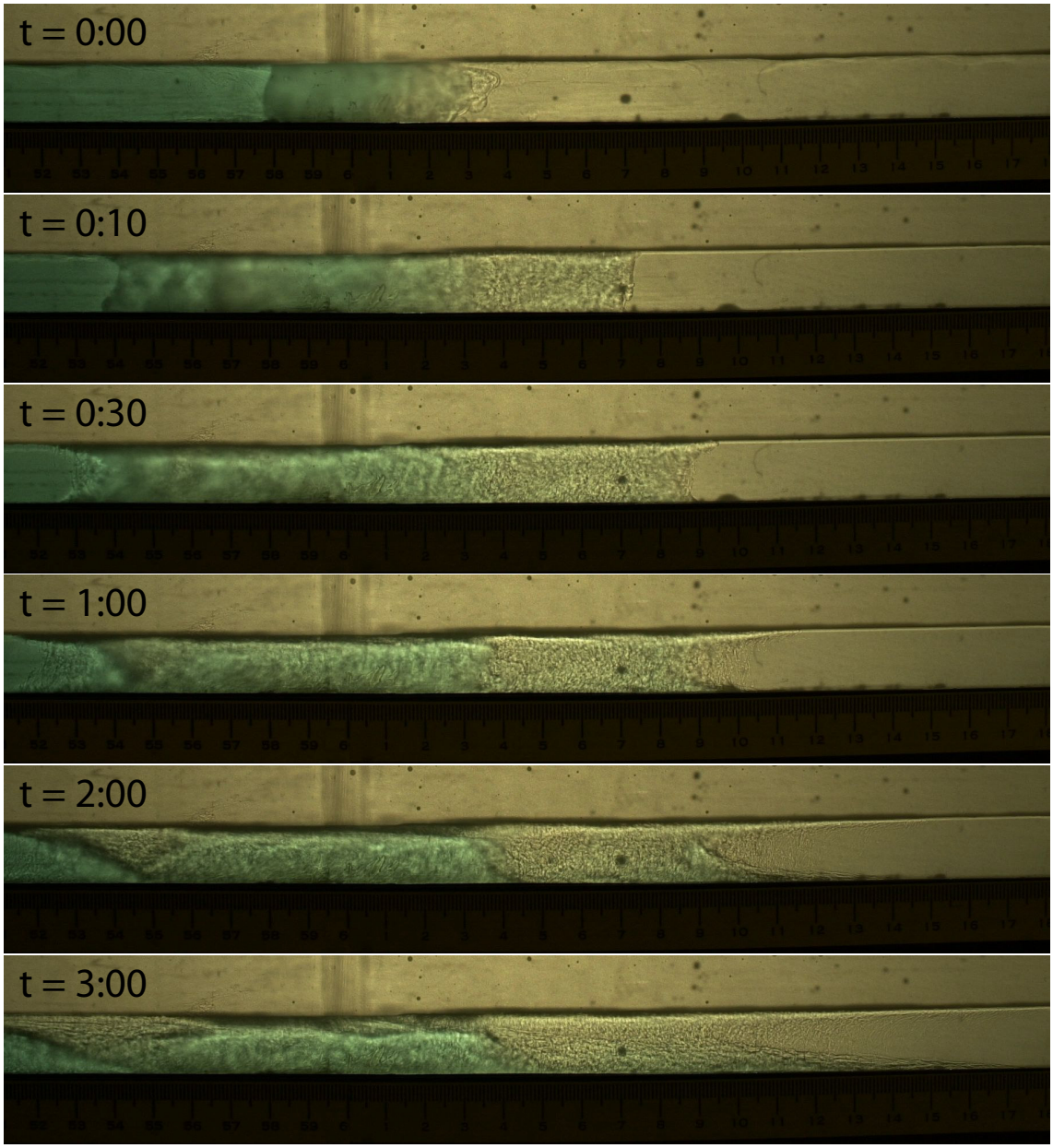


Figure 23: Evolution of a FF gravity current

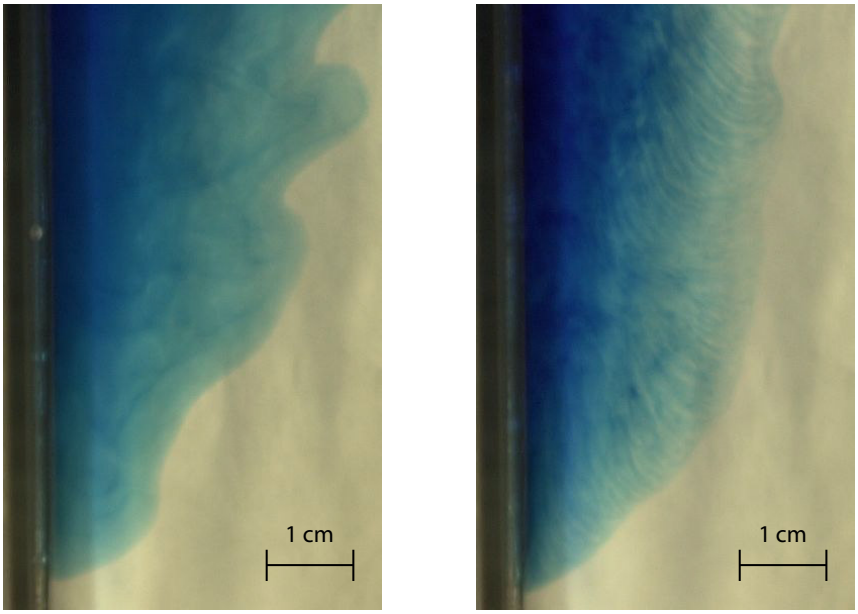


Figure 24: Topview profiles of rotating DF (left) and FF (right) gravity currents with a large density difference and small concentration

and the regions will be not be as well-defined in the presence of convective mixing. But after the fluid settles, this activity would result in the formation of two gravity currents, both a very light current atop and a heavy current along the bottom, as reported in prior experimental studies [13, 14], and is most pronounced when the density differences are very small and solute concentration is large. Photographs illustrating this process are shown in Figure 23. In such cases, the top current is nearly invisible and the bottom current is the dominant structure, making it appear that the light and heavy fluids have switched roles. The bottom current is presumably DF, while the top may be weakly FF. Figure 24 clearly shows the presence of fingers atop of a FF current with large density difference and small concentration, supporting this hypothesis.

6 Conclusions and Future Work

In this study we have attempted to determine the role of double diffusion and rotation on gravity currents by examining their response to various double diffusive profiles, initial density differences, and solute concentrations. Double diffusive currents exhibited a greater head speed than nondiffusive currents that were otherwise identical, with the DF currents being faster than the FF currents, which is consistent with previous studies. We also find that head speeds increase when the density difference and solute concentration between layers become greater. The head speed is therefore much more complicated than for a nondiffusive current; very large density differences in previous studies that were at least an order of magnitude greater than in the present study produced currents that scaled like $\sqrt{\Delta\rho}$, while for the currents presented here, $\Delta\rho$ is clearly increasing with time and the speed

does not scale simply like the initial density difference. This fact, along with the increased sensitivity to concentration, suggest that double diffusive fluxes can play a significant role in frontal propagation of gravity currents. In any case, the head speeds are observed to be independent of rotation, regardless of the relative importance of double diffusion.

Rotation has the greatest effect on the lateral structure of these currents. Boundary currents form in all rotating cases, but the structure is greatly modified by double diffusion. The shear-like crests associated with the nondiffusive rotating currents are no longer present when the density profiles are double diffusive. Wall separation, which seems to be associated with very small density differences in nondiffusive currents, does not seem to occur. In addition, the crossflow spread is much greater when double diffusion is introduced. This spreading, undoubtedly due to the double diffusive fluxes and associated increase of $\Delta\rho$, may very well be responsible for the reduction of frontal instabilities and lack of wall separation.

To better understand the dynamics behind propagation of the fronts and their lateral structure, especially concerning the formation of instabilities and rate of spread, a robust scale analysis that incorporates the density difference and solute concentration into the double diffusive flux is required. Further experiments should also be conducted to establish the sensitivity of the currents to these parameters. Finally, the role of viscosity must be resolved. A simple theoretical treatment of certain idealized flows would help greatly in all of these efforts.

7 Acknowledgements

This project would not have been possible without George Veronis, who not only suggested the topic but also provided continuous guidance in both the laboratory and the subsequent discussions. I will always remember our many conversations together, and I may have even learned to enjoy softball. I am also grateful to Keith Bradley for his time and effort in setting up the more elaborate parts of the experiments, and teaching me how to do the rest. I also want to acknowledge helpful discussions with Karl Helfrich and Jack Whitehead. Finally, I want to thank all of the GFD summer program staff and other fellows for providing me with an incredible educational opportunity and happy memories of my time here.

References

[1] J. J. Stoker, *Water Waves* (Interscience, New York, 1957).

[2] T. B. Benjamin, "Gravity currents and related phenomena," J. Fluid Mech. **31**, 209 (1968).

[3] J. W. Rottman and J. E. Simpson, "Gravity currents produced by instantaneous releases of a heavy fluid in a rectangular channel," J. Fluid Mech. **135**, 95 (1983).

[4] R. W. Griffiths, "Gravity currents in rotating systems," Ann. Rev. Fluid Mech. **18**, 59 (1986).

[5] C. G. Rossby, "On the mutual adjustment of pressure and velocity distributions in certain simple current systems," J. Mar. Res. **1**, 19 (1937).

[6] A. E. Gill, "Adjustment under gravity in a rotating channel," J. Fluid Mech. **77**, 603 (1976).

[7] A. J. Hermann, P. Rhines, and E. R. Johnson, "Nonlinear Rossby adjustment in a channel: beyond Kelvin waves," J. Fluid Mech. **205**, 469 (1989).

[8] K. R. Helfrich, A. C. Kuo, and L. J. Pratt, "Nonlinear Rossby adjustment in a channel," J. Fluid Mech. **390**, 187 (1999).

[9] M. E. Stern, J. A. Whitehead, and B.-L. Hua, "The intrusion of a density current along the coast of a rotating fluid," J. Fluid Mech. **123**, 237 (1982).

[10] J. S. Turner, "Double-diffusive intrusions into a density gradient," J. Geophys. Res. **83**, 2887 (1978).

[11] B. R. Ruddick and J. S. Turner, "The vertical length scale of double-diffusive intrusions," Deep-Sea Res. **26**, 903 (1979).

[12] P. E. Linden, "Salt fingers in steady flow," Geophys. Fluid Dyn. **6**, 1 (1974).

[13] T. Maxworthy, "The dynamics of double-diffusive gravity currents," J. Fluid Mech. **128**, 259 (1983).

[14] J. Yoshida, H. Nagashima, and W. Ma, "A double diffusive lock-exchange flow with small density difference," Fluid Dyn. Res. **2**, 205 (1987).

[15] B. R. Ruddick and T. G. L. Shirtcliffe, "Data for double diffusers: Physical properties of aqueous salt-sugar solutions," Deep-Sea Res. **26A**, 775 (1979).

[16] A. F. Thompson, in *Proceedings of the 2003 Summer Study Program in Geophysical Fluid Dynamics* (Woods Hole Oceanographic Institution, Woods Hole, MA, 2003), pp. 146–165.

[17] M. E. Stern, "Geostrophic fronts, bores, breaking and blocking waves," J. Fluid Mech. **99**, 687 (1980).

[18] M. E. Stern and E. P. Chassignet, "Mechanism of eddy separation from coastal currents," J. Mar. Res. **58**, 269 (2000).

Experiment	Diffusive Tendency	Density Difference	Mean Density
1	Diffusive Layer	0.00046	1.02
2	Diffusive Layer	0.00044	1.02
5	Fingers	0.00046	1.02
6	Nondiffusive	0.00048	1.00
7	Diffusive Layer	0.00053	1.02
8	Fingers	0.00051	1.02
9	Fingers	0.00048	1.02
10	Diffusive Layer	0.00050	1.02
11	Diffusive Layer	0.00011	1.02
12	Diffusive Layer	0.00013	1.04
13	Diffusive Layer	0.00012	1.02
14	Fingers	0.00012	1.04
15	Fingers	0.00014	1.04
16	Fingers	0.00012	1.04
17	Diffusive Layer	0.00013	1.04
53	Fingers	0.00050	1.02
54	Diffusive Layer	0.00050	1.02

Table 3: Table of nonrotating experiments. Experiments 3 and 4 were removed because of operational error.

Experiment	Diffusive Tendency	Density Difference	Mean Density
18	Nondiffusive	0.00049	1.00
19	Fingers	0.00047	1.02
20	Fingers	0.00049	1.02
21	Diffusive Layer	0.00051	1.02
22	Diffusive Layer	0.00046	1.02
23	Fingers	0.00015	1.02
24	Diffusive Layer	0.00002	1.02
25	Diffusive Layer	0.00012	1.02
26	Nondiffusive	0.00010	1.00
27	Diffusive Layer	0.00013	1.02
28	Nondiffusive	0.00050	1.00
29	Diffusive Layer	0.00049	1.02
30	Fingers	0.00049	1.02
31	Nondiffusive	0.00050	1.00
32	Fingers	0.00012	1.04
33	Fingers	0.00012	1.04
34	Diffusive Layer	0.00011	1.04
35	Diffusive Layer	0.00013	1.04
36	Fingers	0.00049	1.02
37	Diffusive Layer	0.00049	1.02
38	Fingers	0.00060	1.04
39	Diffusive Layer	0.00051	1.04
40	Fingers	0.00014	1.04
41	Fingers	0.00011	1.04
42	Diffusive Layer	0.00011	1.04
43	Nondiffusive	0.00039	1.00
44	Nondiffusive	0.00012	1.00
45	Fingers	0.00011	1.04
46	Diffusive Layer	0.00012	1.04
47	Fingers	0.00047	1.04
48	Diffusive Layer	0.00047	1.04
49	Fingers	0.00012	1.02
50	Diffusive Layer	0.00011	1.02
51	Fingers	0.00053	1.02
52	Diffusive Layer	0.00056	1.02

Table 4: Table of rotating experiments. Note that Experiment 24 is only diffusive layer favorable in a marginal sense, since the density difference is so small.

Convection in a Fluid Loop

Eleanor Frajka Williams

Abstract

Sandström's conjecture states that a closed steady circulation can only be maintained in a fluid body if the heating is applied at a lower level than the cooling. In the ocean this has been taken to mean that there can not be a purely buoyancy-driven circulation since the primary sources of heat and salt are applied at the surface. Using the one-dimensional fluid loop of Welander with boundary condition of temperature relaxing to warm and cold at the same geopotential, we find that Sandström's conjecture holds; circulation and horizontal heat transport all vanish as the Rayleigh number $R \longrightarrow \infty$ (the limit relevant for the ocean). However, if we specify fixed temperature flux boundary conditions, then contrary to Sandström's conjecture, circulation does not vanish as $R \longrightarrow \infty$. Thus Sandström's conjecture is sensitive to the particular choice of boundary conditions specified. In the case of the ocean, where salinity and sensible heat also play a role in forcing the system, it is not clear that fixed temperature rather than fixed flux boundary conditions at the surface are appropriate.

This system of horizontal convection in a loop is stable, but becomes chaotic as the heat sources are rotated from horizontal to vertical positions (heating at the bottom and cooling at the top). We determine the orientation of the heating and cooling at which this transition to chaos occurs. We also show how the strength of the circulation responds to different positioning of heating and cooling.

1 Introduction

Oceanographers have long debated whether or not there can exist a purely buoyancy-driven circulation in the ocean. Roots of the discussion trace back at least as far as J.W. Sandström in 1908 who, assuming the ocean circulation worked like a heat engine, concluded that there cannot exist a circulation unless the heating occurs below the cooling, a result sometimes known as Sandström's Theorem. We will call it Sandström's conjecture. The statement has since been formalized to include effects of friction and viscosity, which Sandström did not consider, and can be stated as

> *Sandström's Conjecture:* "If a viscous and diffusive fluid is nonuniformly heated from the same level then in the limit $\kappa \longrightarrow 0$ with ν/κ fixed, the motion in the fluid also disappears," (1).

Oceanographers have inferred from Sandström's conjecture that turbulent mixing rather than buoyancy drives the global conveyor belt (meridional overturning circulation) and that there must be a mechanical energy source to drive the raising of cold water such as wind or tides (2). Then the meridional overturning circulation is just a salt and heat conveyor belt riding on the wind- and

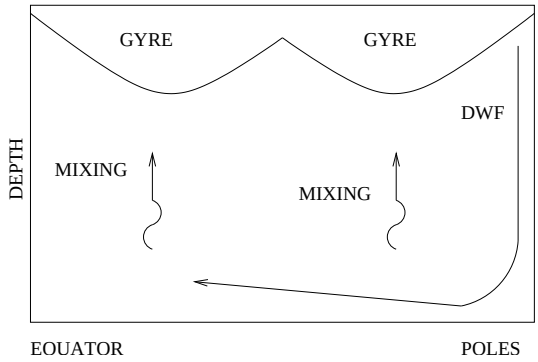
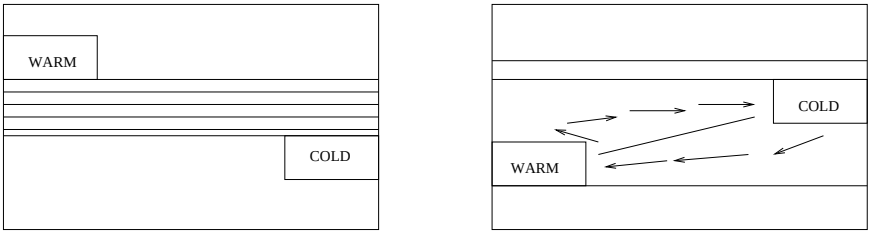


Figure 1: The story: A picture of the global ocean circulation. The wind-driven gyres are near the surface, deep water formation (DWF) at the poles. Mixing due to wind stress and tidal dissipation is required to bring heat down to the abyss to stratify the abyssal ocean. If you remove the mixing, the abyss will fill up with cold water and there will be no circulation.



(a) Type I (b) Type III

Figure 2: Convection in a box with heating and cooling sources placed at different levels. Adapted from Defant (4).

tidally-driven circulation (3). Figure 1 shows the ocean circulation with the wind-driven gyres at the top, with deep convection at the poles closed by interior mixing. In the absence of this mixing, Munk and Wunsch say the ocean would just fill up with cold water and have a very shallow, warm circulation near the surface (2).

Paparella and Young (2002) examined horizontal convection in a 3-D ocean with fixed temperature boundary conditions (5). They showed that dissipation vanishes as diffusivity vanishes using the maximum extremum principle which they contended supports inferences derived from Sandström's conjecture, in particular that, if the ocean were driven purely by surface heating, then one wouldn't see the observed small-scale turbulence. However, their arguments do not eliminate the possibility of an ocean circulation that *is* driven by surface heating.

Using energetics, they showed that, in the steady state, energy is input to potential energy from mixing given by the term $\kappa\langle b_z\rangle$, then converted to kinetic energy by the vertical flux of buoyancy $\langle wb\rangle$ and eventually lost to internal energy through mechanical energy dissipation ε where in all cases $\langle\cdot\rangle$ represents total ocean averages. In steady state, the global conversion terms between

different types of energy are

$$\varepsilon = \langle wb \rangle = \kappa \langle b_z \rangle \, . \tag{1}$$

Averaging mixing $\kappa \langle b_z \rangle$ over z they write,

$$\varepsilon \leq \frac{\kappa}{H} \int (b_{top} - b_{bottom}) dxdy \, . \tag{2}$$

Applying fixed temperature boundary conditions, the *maximum extremum* principle implies that the integral on the right-hand side is bounded by a constant Δb, and they conclude that ε scales with the diffusivity κ. This result, that ε vanishes, is their *anti-turbulence theorem*.

However, using calculus of variations, one can show that globally averaged velocity is bounded by the gradient in velocity times a constant prefactor,

$$|\mathbf{u}|^2 \ \leq \frac{2L^2}{\pi^2} \ |\nabla \mathbf{u}|^2 \quad , \tag{3}$$

where they considered the ocean periodic in both horizontal directions with period L. Since $\varepsilon = \nu \ |\nabla \mathbf{u}|^2$ then for circulation,

$$|\mathbf{u}|^2 \ \leq \frac{2L^2}{\pi^2 H} \frac{\kappa}{\nu} \Delta b \, . \tag{4}$$

Keeping the Prandtl number ν/κ fixed, as $\kappa \longrightarrow 0$, then $|\mathbf{u}|^2 \ \leq$ const and Paparella and Young (2002) neither confirms nor denies Sandström's conjecture.

In this paper, we revisit the idea of a buoyancy driven circulation in a simple 1-D thin fluid loop forced by nonuniform heating and two types of boundary conditions: a temperature relaxation boundary condition and a fixed buoyancy flux boundary condition. In §2 we describe the fluid loop formulation and the energetics and form of solutions in the system. In §3 we examine the behavior of steady state fluid flows in the loop subject to heating and cooling at the same geopotential, as well as implications of these solutions for Sandström's conjecture. In §4, we look at type III convection where heating is applied above cooling, and in §5 we look at type I convection where heating is applied below cooling. In this configuration, we find time dependent solutions including periodic orbits, periodic doubling, and chaos-like behavior. Still, we are able to bound quantities with oceanic import, like mechanical energy dissipation and circulation strength. In section §6, we summarize our conclusions, mention a revision of Sandström's conjecture which is true, but does not have the practical applications of Sandström's original conjecture. Finally, we make a few brief comments about what all this means for the global ocean circulation.

2 Fluid Loop formulation

First introduced by Malkus, Howard and Welander (6; 7; 8), the 1-D fluid loop is one of the simplest systems in which to analyze convection. The single dimension of the system is azimuthal, as a result of the thin loop approximation–that the radial thickness of the loop is vanishingly small, and there is no variation in the radial direction. Here, the system is forced by heating and cooling around the loop, with no net heating or cooling. We will also hold the Prandtl number constant.

In our loop, with radius a, v_θ is the azimuthal velocity, Ω is the angular velocity, and angle θ is defined positive in the counterclockwise direction as shown in Figure 3. The loop is oriented

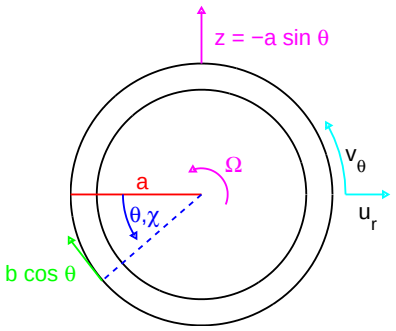


Figure 3: Fluid loop coordinate system and symbol definitions.

vertically so that the z-direction is $a\cos\theta$ and buoyancy acting in the z-direction has component $b\cos\theta$ in the azimuthal direction.

Given this formulation, we may write the velocity of the the fluid as

$$\frac{1}{a}\frac{\partial v_\theta}{\partial \theta} = 0 \,, \qquad v_\theta = a\Omega(t) \,. \tag{5}$$

The buoyancy equation is

$$\frac{\partial b}{\partial t} + \frac{v_\theta}{r}\frac{\partial b}{\partial \theta} = \kappa \left[\frac{\partial^2 b}{\partial r^2} + \frac{1}{r}\frac{\partial b}{\partial r} - \frac{b}{r^2} - \frac{1}{r^2}\frac{\partial^2 b}{\partial \theta^2} \right] \,. \tag{6}$$

We may define our boundary conditions at the inner and outer edges of the thin loop as

$$u_r = 0 \quad \text{at} \quad r_-, r_+ \,, \tag{7}$$

$$\kappa b_r|_{r_-}^{r_+} = \gamma(b^* - b) + F(\theta) \,, \tag{8}$$

where the sources are

$$b^* = g\cos(\theta + \chi) \,, \tag{9}$$

$$F(\theta) = F_0 \cos(\theta + \chi) \,. \tag{10}$$

with zero azimuthally averaged forcing, necessary to be able to consider steady states of the system. In all cases below, we will consider $\chi \in [-\pi, \pi]$. Integrating the buoyancy equation (6) in the radial direction and applying boundary condition (8) gives,

$$\frac{\partial b}{\partial t} + \Omega\frac{\partial b}{\partial \theta} = \gamma(b^* - b) + F(\theta) \,. \tag{11}$$

The momentum equation in the radial direction is

$$\frac{\partial u_r}{\partial t} + u_r\frac{\partial u_r}{\partial r} + \frac{v_\theta}{r}\frac{\partial u_r}{\partial \theta} - \frac{v_\theta^2}{r} = -\frac{1}{\rho}\frac{\partial p}{\partial r} + \eta \left[\frac{\partial^2 u_r}{\partial r^2} + \frac{1}{r}\frac{\partial u_r}{\partial r} - \frac{u_r}{r^2} + \frac{1}{r^2}\frac{\partial^2 u_r}{\partial \theta^2} - \frac{2}{r^2}\frac{\partial v_\theta}{\partial \theta} \right] \,, \tag{12}$$

and for a very thin loop, viscosity dominates in the radial direction, then

$$\frac{\partial^2 u_r}{\partial r^2} \approx 0 \,, \tag{13}$$

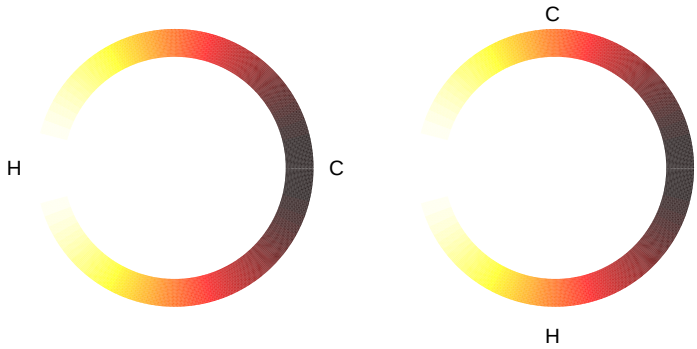


Figure 4: (a) Horizontal convection in the loop ($\chi = 0$). The location of heating maybe thought of as the equator in the ocean, and the location of cooling as the poles. (b) Rayleigh-Bénard convection in the loop ($\chi = 90$). White shading indicates heating; dark shading indicates cooling.

and using the boundary condition (7),

$$u_r = 0 \; . \tag{14}$$

Then the momentum equation in the azimuthal direction becomes

$$\frac{\partial v_\theta}{\partial t} + \frac{v_\theta}{r}\frac{\partial v_\theta}{\partial \theta} = -\frac{1}{r\rho}\frac{\partial p}{\partial \theta} + \eta \quad \frac{\partial^2 v_\theta}{\partial r^2} + \frac{1}{r}\frac{\partial v_\theta}{\partial r} - \frac{v_\theta}{r^2} - \frac{1}{r^2}\frac{\partial^2 v_\theta}{\partial \theta^2} \quad + b\cos\theta \; . \tag{15}$$

From equation (5), we have no variations in velocity or pressure in the azimuthal direction, then integrating in r and θ we have

$$\frac{\partial \Omega}{\partial t} = \frac{1}{2\pi a}\int b\cos\theta d\theta - \eta\Omega \; . \tag{16}$$

Given these sources, when $\chi = 0$ we have heating and cooling applied at the same geopotential level as shown in Figure 2. When $\chi < 0$, heating is applied above cooling as shown in Figure 2. When $\chi > 0$, we have heating applied below cooling in the loop. We will consider these three cases of convection in §3-5.

2.1 Energetics

In this system, kinetic and potential energy are given by

$$KE = \frac{1}{2}a^2\Omega^2 \; , \quad PE = a\cos\theta b \; . \tag{17}$$

By taking $v_\theta \cdot (\text{momentum})$ and integrating around the loop, we can write an evolution equation for kinetic energy as

$$\langle KE \rangle_t = -a\Omega \langle b\cos\theta \rangle - \eta a^2\Omega^2 \; , \tag{18}$$

where $\langle\cdot\rangle = \frac{1}{2\pi}\int \cdot d\theta$. On the right hand side, the first term is the vertical buoyancy flux, the conversion term between kinetic and potential energy, analogous to $\langle wb\rangle$ in the cartesian coordinate system, and the second term is the mechanical energy dissipation ε.

The evolution equation for potential energy is derived by deriving in time $a\langle b\cos\theta\rangle$,

$$\langle PE\rangle_t = a\Omega\langle b\cos\theta\rangle + a\gamma\langle b\sin\theta\rangle + \frac{a}{2}(\gamma g + F_0)\sin\chi \,. \tag{19}$$

The first term in (19) is again the conversion term between kinetic and potential energy. The second term is the source of potential energy through mixing, analogous to $\langle\kappa b_z\rangle$ in the cartesian coordinate system, and the third term representing sources of potential energy through external heating and cooling.

The boundary conditions add no net buoyancy to the system, so the long-time averages of the energy equations (18) and (19) vanish, and we can write

$$\eta a^2\overline{\Omega^2} = a\overline{\Omega\langle b\cos\theta\rangle} = a\gamma\overline{\langle b\sin\theta\rangle} + \frac{a}{2}(\gamma g + F_0)\sin\chi \,. \tag{20}$$

where $\bar{\cdot} = \lim_{T\longrightarrow\infty}\frac{1}{T}\int_0^T \cdot dt$, the long-time average.

Horizontal Convection ($\chi = 0$)

For horizontal convection, equation (20) becomes

$$\eta a^2\overline{\Omega^2} = a\overline{\Omega\langle b\cos\theta\rangle} = a\gamma\overline{\langle b\sin\theta\rangle} \,,$$

which is analogous to Paparella and Young's statement that

$$\varepsilon = \langle wb\rangle = \kappa\langle b_z\rangle \,.$$

However, the same arguments that they used to bound dissipation by a constant times diffusivity do not hold in this case. In particular, the *maximum extremum* principle no longer applies and one cannont bound $\langle b_z\rangle$ nor $a\gamma\overline{\langle b\sin\theta\rangle}$ by a constant, so it is not immediately obvious that the *anti-turbulence theorem* should hold.

Heating above cooling ($\chi < 0$)

In the case of heating above cooling, equation (20) becomes

$$\eta a^2\overline{\Omega^2} = a\overline{\Omega\langle b\cos\theta\rangle} = a\gamma\overline{\langle b\sin\theta\rangle} - \frac{a}{2}(\gamma g + F_0)|\sin\chi| \,.$$

Added buoyancy through either the relaxation temperature or fixed flux boundary condition represents a loss of potential energy to drive the circulation.

Heating below cooling ($\chi > 0$)

For heating below cooling, we have

$$\eta a^2\overline{\Omega^2} = a\overline{\Omega\langle b\cos\theta\rangle} = a\gamma\overline{\langle b\sin\theta\rangle} + \frac{a}{2}(\gamma g + F_0)|\sin\chi| \,.$$

and the buoyancy sources represent a source of potential energy available to drive circulation.

2.2 Solutions for b

Since we are considering solutions in a fluid loop, the buoyancy solution must necessarily be periodic. We can construct the general solution of buoyancy in our equations as an infinite sum of sine and cosine modes,

$$b(t,\theta) = \sum_{n=1}^{\infty} S_n(t)\sin(n\theta) + C_n(t)\cos(n\theta)\ . \tag{21}$$

where S_n and C_n are the amplitudes of the nth mode of buoyancy. Using this form of the solution, equations (16) and (11) become

$$\frac{\partial \Omega}{\partial t} = -\frac{1}{2\pi a}\int C_1(t)\cos^2\theta d\theta - \eta\Omega\ , \tag{22}$$

$$\frac{\partial S_n}{\partial t} = \Omega C_n - \gamma S_n + (\gamma g + F_0)\sin\chi\ , \tag{23}$$

$$\frac{\partial C_n}{\partial t} = -\Omega S_n - \gamma C_n + (\gamma g + F_0)\cos\chi\ . \tag{24}$$

From these equations we can see that only the mode one ($n = 1$) solutions for buoyancy are coupled to equation (22), so the full dynamics of the system are captured using only the mode one solution, irrespective of the form of the forcings, b^* and $F(\theta)$. So dropping the subscript n, we use

$$b(t,\theta) = S(t)\sin\theta + C(t)\cos\theta \tag{25}$$

as our buoyancy solution, and

$$\Omega_t = -\frac{C}{2\pi a}\int \cos^2\theta d\theta - \eta\Omega = -\frac{C}{2a} - \eta\Omega\ , \tag{26}$$

$$S_t = \Omega C - \gamma S + (\gamma g + F_0)\sin\chi\ , \tag{27}$$

$$C_t = -\Omega S - \gamma C + (\gamma g + F_0)\cos\chi\ . \tag{28}$$

are the equations governing our system.

2.3 Nondimensional equations

We rescale the system with the following nondimensionalization,

$$\begin{array}{llll} \hat{\Omega} = \Omega/\gamma, & \hat{S} = \frac{S}{2a\gamma\eta}, & R = \frac{g}{2a\gamma\eta}, & \sigma = \eta/\gamma, \\ \hat{t} = \gamma t, & \hat{C} = \frac{C}{2a\gamma\eta}, & R_F = \frac{F_0}{2a\gamma^2\eta}. & \end{array}$$

where the $\hat{}$ denotes the nondimensional form of a variable. The Rayleigh number is R in our system, and we write an analogous Rayleigh number R_F for the fixed buoyancy flux term.

We rename the variables

$$\begin{array}{ll} X = \hat{\Omega}, & Z = -\hat{C} + r\sin\chi, \\ Y = \hat{S}, & r = R + R_F\ , \end{array}$$

which gives us a new system of equations

$$\dot{X} = \sigma(Y - X) \ , \tag{29}$$

$$\dot{Y} = r\sin\chi X - Y - XZ + r\cos\chi \ , \tag{30}$$

$$\dot{Z} = XY - Z \ . \tag{31}$$

When $\chi = 90$, this is these are the Lorenz equations,[1] where heating and cooling are applied at the bottom and top of the loop respectively. In this system, X is our nondimensional angular velocity, Y is the sine buoyancy mode amplitude, and Z is the cosine buoyancy mode amplitude shifted and flipped. More physical quantities which we will consider are the nondimensional dissipation $\hat{\varepsilon}$, buoyancy flux $\widehat{\langle wb \rangle}$, and mixing $\widehat{\langle \kappa b_z \rangle}$,

$$\hat{\varepsilon} = \sigma X^2 \ , \quad \widehat{\langle wb \rangle} = \sigma XY \ , \quad \widehat{\langle \kappa b_z \rangle} = \sigma Z \ . \tag{33}$$

We will also consider horizontal heat transport, $H_T = \langle v_\theta b \rangle = -a\Omega S/2$. In nondimensional form,

$$\widehat{H_T} = -\sigma(XZ + r\sin\chi X) \ . \tag{34}$$

In the ocean, heating and cooling are applied at the ocean surface which is equivalent to being applied at the same geopotential in the loop, at $\chi = 0$. We will examine the steady state solutions for buoyancy and circulation, their stability, and some of the physical quantities in equation (33) as functions of the Rayleigh numbers defined in our nondimensionalization (§2.3). The goal is to determine the behavior of the solution in the diffusive and nondiffusive limits (small r and large r, respectively).

We will then proceed to consider convection where the heating is applied from above (at $\chi < 0$), the stability of these steady state solutions and their implications for dissipation and circulation. Finally we will consider the case of heating applied below cooling, which is the atmospheric analog since the atmosphere is relatively transparent to the short-wave radiation from the sun, and is heated by long-wave radiation from the Earth. We will again attempt to say something about the behavior of dissipation, circulation and horizontal heat transport as functions of the Rayleigh numbers.

The system of equations (29)-(31) has steady solutions, given by the fixed points (X_0, Y_0, Z_0) satisfying

$$\begin{aligned} Y_0 &= X_0 \ , \\ X_0^3 + (1 - r\sin\chi)X_0 &= r\cos\chi \ , \\ Z_0 &= X_0^2 \ . \end{aligned} \tag{35}$$

If the fixed points are globally stable, then the system has no intrinsic time dependence. If the fixed points are not stable, the solutions of X, Y, and Z to the full equations (29)-(31) can develop time dependence.

[1] The Lorenz system is

$$\begin{aligned} \dot{X} &= \sigma(Y - X) \\ \dot{Y} &= rX - Y - XZ \\ \dot{Z} &= XY - bZ \end{aligned} \tag{32}$$

where b is the aspect ratio of convective cells, typically taken to be 8/3 in the atmosphere, but for a circular fluid loop, $b = 1$ as in our case. Tritton showed that the Lorenz equations could be derived for the loop (9).

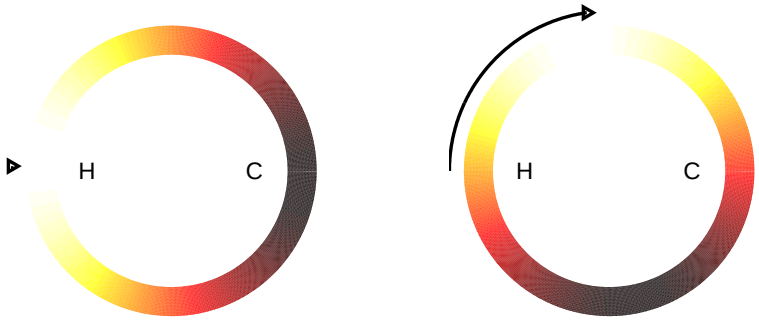


Figure 5: Steady state solutions for horizontal convection $\chi = 0$ with (a) $r = 0.1$ and (b) $r = 100$. "H" and "C" indicate locations of maximal applied heating and cooling and may be thought of as equator and polar regions in the ocean. Colors indicate temperature in the steady state solutions (white hot, dark cold), and arrows indicate the direction and relative magnitudes of circulation for the two cases, small and large r.

3 Horizontal Convection ($\chi = 0$)

For the case of heating and cooling at the same level, also known as horizontal convection, we have $\chi = 0$, for which the fi xed point equation (35) becomes

$$X_0^3 + X_0 = r \ . \tag{36}$$

In the approximation of large and small r, we have the balances

$$X_0^3 \approx r \quad \text{for } r \text{ large} \ , \tag{37}$$

$$X_0 \approx r \quad \text{for } r \text{ small} \ , \tag{38}$$

which give the scalings for the nondimensional, steady state circulation X_0.

Exact solutions are shown in Figure 5. The locations of maximum applied heating and cooling are indicated by the letters "H" and "C", while the colors in the loop show the steady state pattern of temperature in the loop. The arrows show the magnitude and direction of the resultant circulation. In terms of the buoyancy solution in equation (25), the applied forcing for horizontal convection is a cosine mode. For small r (large diffusivity), the resulting steady state pattern of warm and cool water in the loop is dominated by the cosine mode. The pattern is displaced slightly, however, with warm water above the location of heating since the circulation displaces some water, and the diffusion (except in the $r = 0$ limit) does not fully relax the temperature pattern to the applied heating. The circulation is small, but nonzero, resulting from the applied torque of the nonuniform heating, balanced by frictional drag. For large r the pattern of buoyancy is dominated by the sine mode, 90 out-of-phase, and the circulation carries the heat towards the top of the loop and a position of smaller potential energy in the system. The patterns of temperature in the steady state shows the dominance of the forces on the either diffusive, relaxing the temperature pattern to the applied heating, or convective, reducing the torque by moving the light water to the top and the heavy water to the bottom.

3.1 Stability of solutions ($\chi = 0$)

Numerical experiments suggested that the steady solutions were reached for any arbitrary condition. Global stability of a solution means that from any point will move in time towards the steady solution and won't stop till it gets there. This can be shown by constructing a Lyaponov functional of a perturbation (not necessarily small) around the fixed point, then show that the derivative of the Lyaponov functional is negative. Mathematically speaking, we find a function $V(\delta\mathbf{x})$ where $\mathbf{X} = (X,Y,Z)$ and $\mathbf{X} = \mathbf{X_0} + \delta\mathbf{x}$, such that $\dot{V} < 0 \;\forall\; \delta\mathbf{x} \neq \mathbf{0}$, and $\dot{V} = 0$ only for $\delta\mathbf{x} = \mathbf{0}$. We construct a Lyaponov functional for this system, not necessarily unique, as

$$V(\delta x, \delta y, \delta z) = \frac{X_0^2}{\sigma} \quad \delta x^2 + \delta y^2 + \delta z^2 \; . \tag{39}$$

The derivative of V may be rearranged

$$\frac{1}{2}\dot{V} = - \quad \frac{X_0^2}{2} \quad \delta x^2 - \delta y^2 - \frac{1}{2}\delta z^2 - \frac{1}{2}(\delta z - \delta x X_0)^2 \; , \tag{40}$$

from which we can see that

$$\dot{V} \leq 0 \; , \tag{41}$$

so solutions to equation (36) are globally stable.

3.2 Circulation, dissipation and horizontal heat transport ($\chi = 0$)

In the steady state, X_0 is the nondimensional circulation, and nondimensional dissipation and horizontal heat transport are given in the equations in (33). Dissipation and horizontal heat transport scale with r in the large and small limits as

$$\hat{\varepsilon} \approx \sigma r^2 \; , \quad \widehat{H_T} \approx -\sigma r^3 \quad \text{for } r \text{ small} \; , \tag{42}$$

$$\hat{\varepsilon} \approx \sigma r^{2/3} \; , \quad \widehat{H_T} \approx -\sigma r \quad \text{for } r \text{ large} \; . \tag{43}$$

Using the exact solutions of the fixed point equation, the circulation, dissipation and horizontal heat transport shown in Figure 6a-c, and match these scalings in each limit.

In the dimensional form, these quantities are dependent on the choice of boundary conditions used, specified as sources in equation (11), or for the nondimensional fixed point equation as,

$$X_0^3 + X_0 = R \; , \tag{44}$$

$$X_0^3 + X_0 = R_F \; . \tag{45}$$

These two equations are equivalent to setting either $F(\theta) = 0$, or $b^* = 0$ in the dimensional equations (23), which we called restoring-like or fixed flux-like boundary conditions, respectively. The behavior of the dimensional quantities circulation, dissipation and horizontal heat transport

$$\Omega^2 = \gamma^2 X_0^2 \; , \quad \varepsilon = \eta a^2 \Omega^2 \; , \quad H_T = -\sigma a^2 \Omega^3 \; . \tag{46}$$

will be considered as a function of the relaxation timescale, or equivalently diffusivity, γ.

Restoring-like boundary conditions

For restoring-like boundary conditions, we use equation (44) to determine dimensional scalings. In the steady state, dimensional circulation given by $\Omega = \gamma X_0$ scales as

$$\Omega \approx \frac{g}{2a\sigma} \longrightarrow 0 \text{ as } \gamma \longrightarrow \infty \,, \tag{47}$$

$$\Omega \approx \left(\frac{g}{2a\sigma}\right)^{1/3} \gamma^{1/3} \longrightarrow 0 \text{ as } \gamma \longrightarrow 0 \,. \tag{48}$$

So in both the nondiffusive and diffusive limits, circulation vanishes. For dissipation,

$$\varepsilon \approx \frac{g^2}{4\sigma}\gamma^{-1} \longrightarrow 0 \text{ as } \gamma \longrightarrow \infty \,, \tag{49}$$

$$\varepsilon \approx \left(\frac{g^2 a^4 \sigma}{4}\right)^{1/3} \gamma^{5/3} \longrightarrow 0 \text{ as } \gamma \longrightarrow 0 \,. \tag{50}$$

Again, in both limits, dissipation vanishes. Lastly, horizontal heat transport in the steady state,

$$H_T \approx -\frac{g^3}{8a\sigma^2}\gamma^{-3} \longrightarrow 0 \text{ as } \gamma \longrightarrow \infty \,, \tag{51}$$

$$H_T \approx -\frac{ga}{2}\gamma \longrightarrow 0 \text{ as } \gamma \longrightarrow 0 \,, \tag{52}$$

vanishes in both limits. These scalings are shown in Figure 6d-f as indicated by the blue lines.

Fixed flux-like boundary conditions

In the steady state, dimensional circulation given by $\Omega = \gamma X_0$ scales as

$$\Omega \approx \frac{F_0}{2a\sigma}\gamma^{-2} \longrightarrow 0 \text{ as } \gamma \longrightarrow \infty \,, \tag{53}$$

$$\Omega \approx \left(\frac{F_0}{2a\sigma}\right)^{1/3} \longrightarrow \text{ const as } \gamma \longrightarrow 0 \,. \tag{54}$$

So in the diffusive limit, circulation vanishes, but in the nondiffusive limit (as $\gamma \longrightarrow 0$), circulation remains constant. For dissipation,

$$\varepsilon \approx \frac{F_0^2}{4\sigma}\gamma^{-3} \longrightarrow 0 \text{ as } \gamma \longrightarrow \infty \,, \tag{55}$$

$$\varepsilon \approx \left(F_0^2 a^4 \sigma\right)^{1/3} \gamma \longrightarrow 0 \text{ as } \gamma \longrightarrow 0 \,. \tag{56}$$

In both limits, dissipation vanishes. Horizontal heat transport in the steady state,

$$H_T \approx -\frac{F_0^3}{8a\sigma^2}\gamma^{-6} \longrightarrow 0 \text{ as } \gamma \longrightarrow \infty \,, \tag{57}$$

$$H_T \approx -\frac{F_0 a}{2} \longrightarrow \text{ const as } \gamma \longrightarrow 0 \,, \tag{58}$$

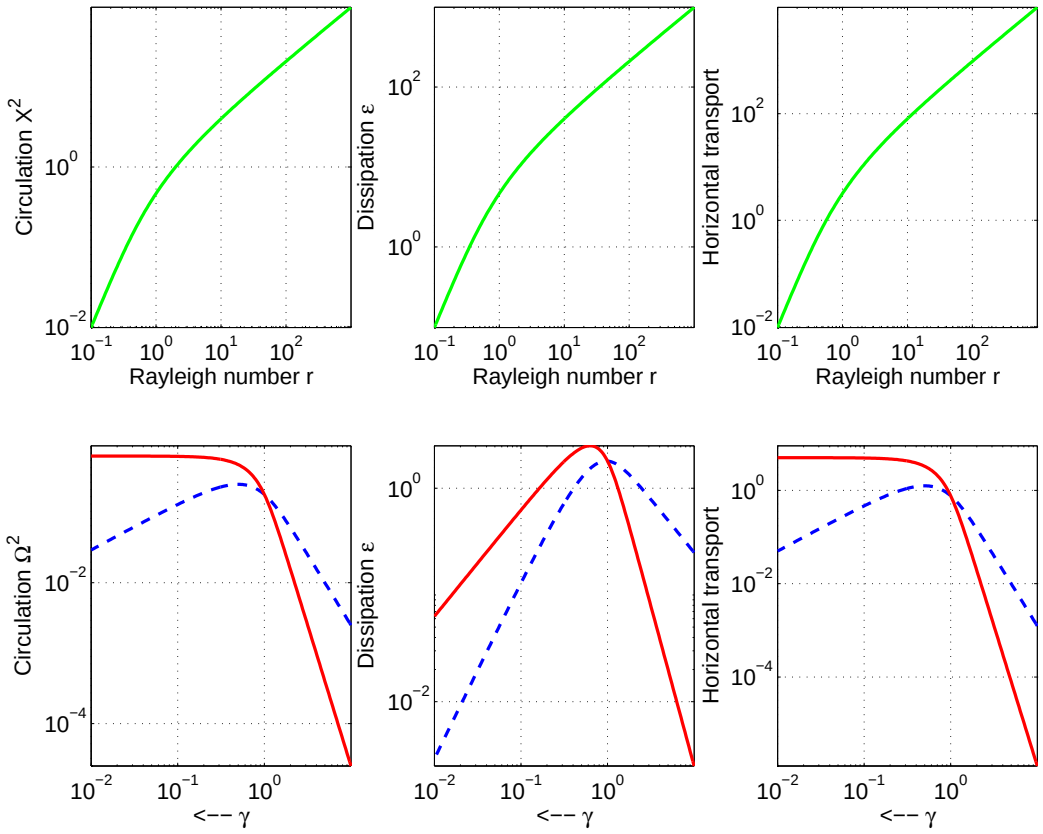


Figure 6: (a)-(c) Nondimensional scaling of circulation, dissipation and horizontal heat transport as functions of the Rayleigh number. (d)-(f) Dimensional scaling of the same, as functions of the relaxation timescale γ. Red solid lines indicate solutions using only fixed buoyancy flux boundary conditions. Blue dashed lines indicate solutions using only the relaxation temperature boundary conditions.

vanishes in the diffusive limit, but is a constant in the nondiffusive limit. These scalings are shown in Figure 6d-f as indicated by the red lines.

In summary, In both cases, dissipation vanishes, in agreement with Paparella and Young's *anti-turbulence theorem*, and for the restoring-like boundary conditions, circulation, and horizontal heat transport vanish in the nondiffusive limit (the oceanic case). However for fixed flux-like boundary conditions, circulation and horizontal heat transport do not vanish.

3.3 Implications for Sandström's conjecture ($\chi = 0$)

1. In the diffusive limit $\gamma \longrightarrow \infty$, the temperature in the steady solution is relaxed to match the applied forcing while, in the nondiffusive limit $\gamma \longrightarrow 0$, the solution has warm water at the top of the loop and cold water at the bottom, towards the limit of minimum potential energy.

2. In the relaxation temperature case ($F(\theta) = 0$, $b^* \neq 0$), dissipation, circulation and horizontal heat transport all vanish in the nondiffusive limit ($\gamma \longrightarrow 0$). This is in accordance with

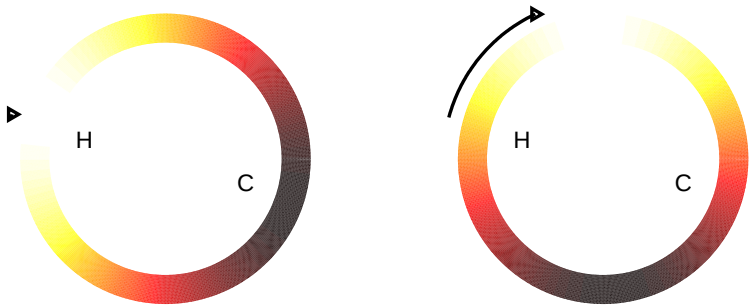


Figure 7: Steady state solutions for heating above cooling $\chi = -15$ with (a) $r = 0.1$ and (b) $r = 100$. "H" and "C" indicate locations of maximal applied heating and cooling. Colors indicate temperature in the steady state solutions, and arrows indicate the direction and relative magnitudes of circulation for the two cases, small and large r.

Sandström's conjecture, that horizontally applied heating cannot drive a circulation.

3. In the fixed buoyancy flux case ($b^* = 0$, $F(\theta) \neq 0$), dissipation still goes to zero in the non-diffusive limit, but circulation and horizontal heat transport saturate at a constant. When diffusion does not carry the applied heating to the location of applied cooling, a circulation must exist to maintain the applied flux. Sandstr öm's conjecture no longer applies given fixed flux boundary conditions. Though heating is still applied horizontally, there is a substantial circulation in the limit of small diffusion.

4 Heating Above Cooling ($\chi < 0$)

For the case of heating above cooling ($\chi < 0$), the fixed point equation (35) is

$$X_0^3 + (1 - r\sin\chi)X_0 = r\cos\chi \,,$$

with $r\sin\chi < 0$. In the approximation of large and small r, we have the balances

$$X_0^3 \approx 1/\tan\chi \quad \text{for } r \text{ large} \,, \tag{59}$$

$$X_0 \approx r\cos\chi \quad \text{for } r \text{ small} \,, \tag{60}$$

which give the scalings for the nondimensional, steady state circulation X_0. In particular, in the large Rayleigh number limit, X_0 scales as a constant, dependent on the angle of applied heating.

The solutions to this cubic polynomial are in Figure 7. Again at low Rayleigh number the circulation is very small and the pattern of heating is very similar to the applied pattern of heating. The circulation results from the balance between applied torque and drag. For large r, the circulation is faster, but not as fast as for horizontal convection at the same Rayleigh number.

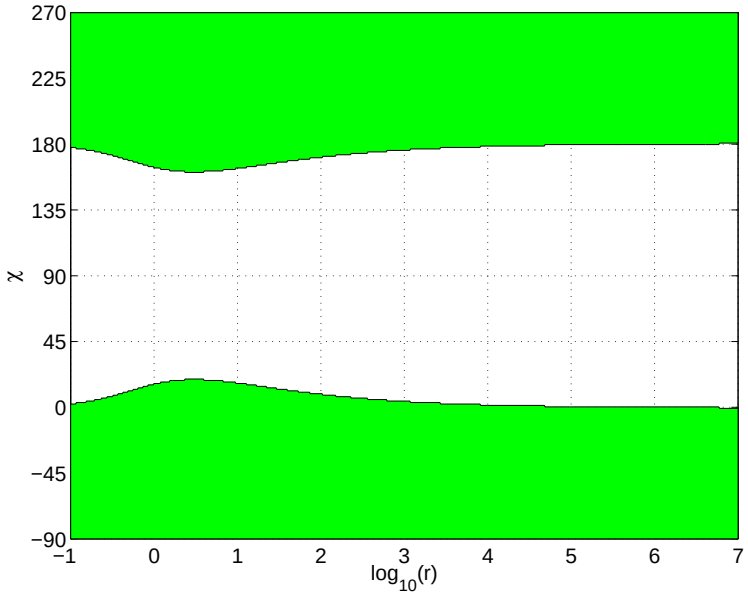


Figure 8: A diagram of the global stability of fi xed points X_0 in $\chi - r$ space given by the condition that $r\sin\chi \leq X_0^2/2$ where X_0 depends in χ and r. The area in green is globally stable.

4.1 Stability of Solutions ($\chi < 0$)

The stability of solutions for heating above cooling is also globally stable, and can be shown using the same Lyaponov functional as for horizontal convection.

$$V(\delta x, \delta y, \delta z) = \frac{X_0^2 - r\sin\chi}{\sigma}\ \delta x^2 + \delta y^2 + \delta z^2 \ . \tag{61}$$

The derivative of V may be rearranged

$$\frac{1}{2}\dot{V} = -(\frac{X_0^2}{2} - r\sin\chi)\delta x^2 - \delta y^2 - \frac{1}{2}\delta z^2 - \frac{1}{2}(\delta z - \delta x X_0)^2 \ , \tag{62}$$

from which we can see that

$$\dot{V} \leq 0 \tag{63}$$

as long as

$$r\sin\chi \leq \frac{1}{2}X_0^2 \ . \tag{64}$$

For all $\chi < 0$, $r\sin\chi \leq 0 \leq X_0^2/2$ and the condition is met, so the solutions are globally stable. The regions in $\chi - r$ space which are globally stable are shown in Figure 8.

4.2 Circulation, dissipation and horizontal heat transport ($\chi < 0$)

For $\chi < 0$, dissipation and horizontal heat transport scale with r in the large and small limits as

$$\hat{\varepsilon} \approx \sigma r^2 \cos^2\chi \ , \quad \widehat{H_T} \approx -\sigma r^3 \cos^3\chi \quad \text{for } r \text{ small} \ , \tag{65}$$

$$\hat{\varepsilon} \approx \sigma \cot^2\chi \ , \quad \widehat{H_T} \approx -\sigma \cot^3\chi \quad \text{for } r \text{ large} \ . \tag{66}$$

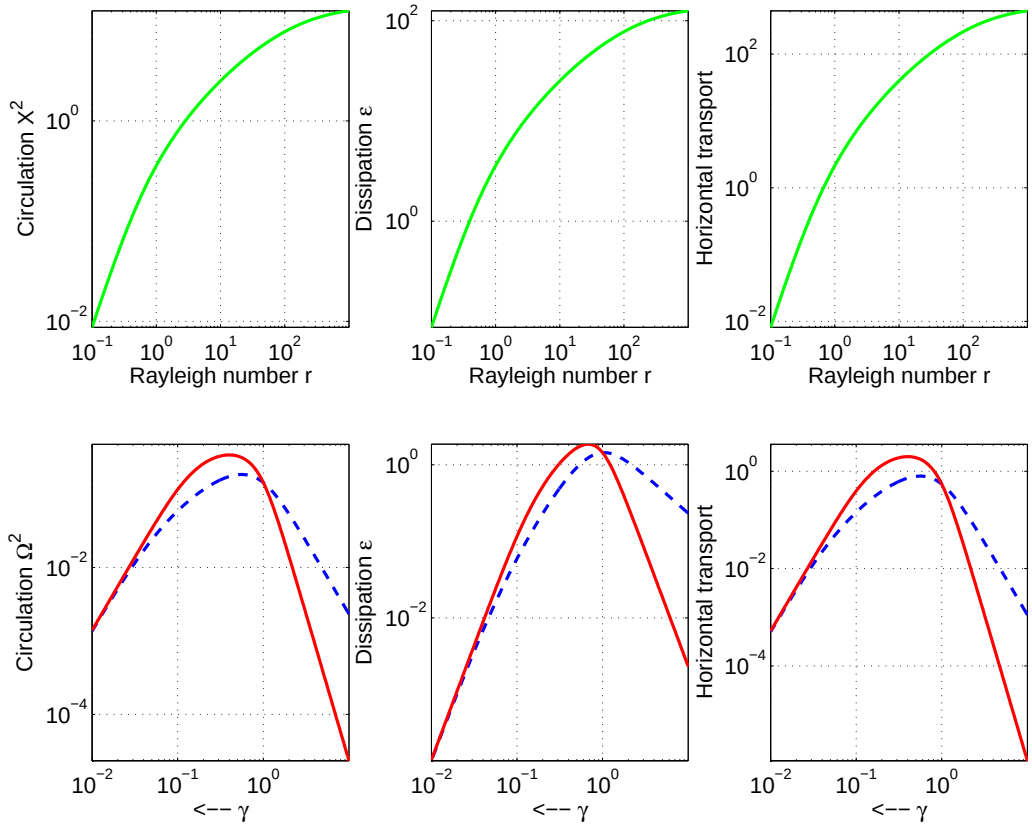


Figure 9: (a)-(c) Nondimensional scaling of circulation, dissipation and horizontal heat transport as functions of the Rayleigh number for heating above cooling ($\chi = -15°$). (d)-(f) Dimensional scaling of the same, as functions of the relaxation timescale γ. Red indicates solutions using only fixed buoyancy flux boundary conditions. Blue indicates solutions using only the relaxation temperature boundary conditions.

These scalings are shown in Figure 9a-c. All three nondimensional quantities saturate at a constant (in r, not in χ) for large r.

In the dimensional form with the quanitites given by the equations in 46, circulation, dissipation and horizontal convection scale as shown in Figure 9d-f. For heating above cooling, with either type of boundary conditions, all dimensional quantities vanish in the diffusive and nondiffusive limits. There exists a circulation for small γ, but it is very weak.

5 Heating Below Cooling ($\chi > 0$)

In this section we will consider heating below cooling, starting first with the special case of heating directly at the bottom of the loop. This case is symmetric across the z-axis and as a result has different characteristics than the other cases for $\chi > 0$. The fixed point equation is, as for $\chi < 0$,

$$X_0^3 + (1 - r\sin\chi)X_0 = r\cos\chi\,, \tag{67}$$

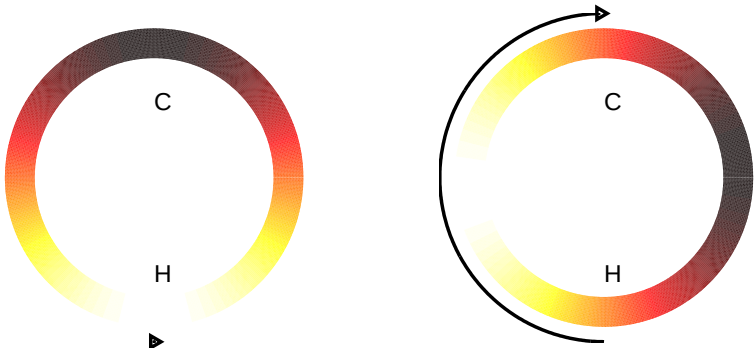


Figure 10: Fixed point solutions for convection with $\chi = 90$ with (a) $r = 0.1$ and (b) $r = 100$. "H" and "C" indicate locations of heating and cooling along the loop. The colors indicate the temperature pattern for a fixed point in the system. Arrows indicate direction and relative magnitude of circulation.

but we will find that for certain parameter ranges of χ and r, the fixed point solutions are not stable, so the solution to the full equations is time dependent.

5.1 Rayleigh-Bénard convection ($\chi = 90$)

We will first consider the special case of heating at the bottom of the loop ($\chi = 90$), called Rayleigh-Bénard convection for which equations (29)-(31) are the Lorenz equations. For this case, the fixed point equation is

$$X_0^3 + (1-r)X_0 = 0\ , \tag{68}$$

with solutions $X_0 = 0$ for all r, and $X_0 = \pm\sqrt{r-1}$ for $r > 1$. This first solution has warm water at the bottom with the heat being carried to the top by diffusion alone. Because the heat is applied symmetrically about the z-axis, there is no applied torque. As diffusion is decreased (r is increased), the $X_0 = 0$ solution becomes unstable (at $r = 1$) and two stable solutions with circulation either clockwise or counterclockwise appear. This is called a pitchfork bifurcation. Two fixed point solutions for small and large r are shown in Figure 10.

For $r = \sigma\frac{\sigma+4}{\sigma-2}$, these two stable solutions become unstable and both simultaneously undergo a subcritical Hopf bifurcation, where the stable solution becomes an unstable limit cycle. The locations of the stable and unstable fixed points, and a diagram of the bifurcation is shown in Figure 11. The pitchfork bifurcation appears as it is called, and the Hopf bifurcations are the two simultaneous from stable to unstable fixed point. More details on this system the Hopf bifurcations in the Lorenz system can be found in Drazin (10).

5.2 General Heating Below Cooling ($\chi \geq 0$)

For the general case of heating below cooling ($\chi > 0$), there are two qualitatively different behaviors that may arise in our system, depending on how close χ is to the symmetric case of $\chi = 90$. For all $\chi > 0$, there is no fixed point solution without circulation since the asymmetrically applied heating creates a torque on the system which, even in cases of large but finite diffusivity (small r), creates a nonzero circulation balancing frictional drag. For larger r, this near zero circulation increases

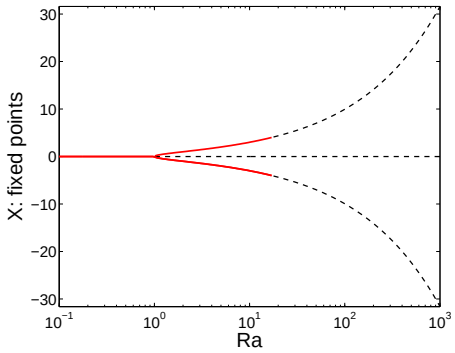


Figure 11: Bifurcation diagram for convection with $\chi = 90$. Solid lines indicate stable fi xed points; dotted lines indicate unstable fi xed points. There is a pitchfork bifurcation at $r = \sim 10$ and two Hopf bifurcations at $r = \sigma(\sigma+4)/(\sigma-2)$.

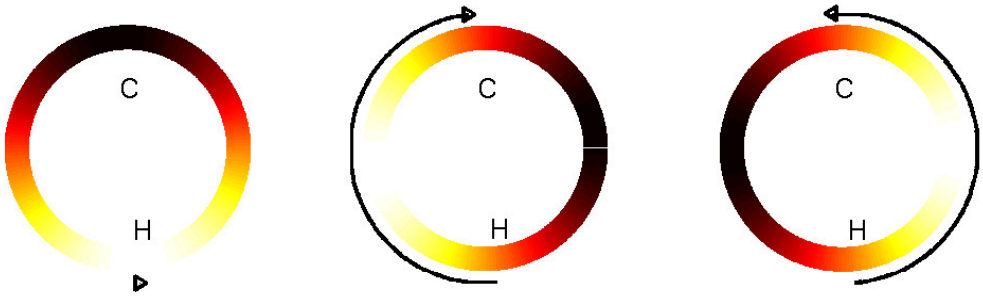


Figure 12: Fixed point solutions for $\chi = 95$. "H" and "C" indicate locations of heating and cooling along the loop. The colors indicate the temperature pattern for a fi xed point in the system. Arrows indicate direction and relative magnitude of circulation.

in magnitude, and two new fi xed points appear out-of-nowhere. The appearance of two new fi xed point solutions from nowhere is called a saddle-node bifurcation. For χ near 90, one of the new fi xed point solutions is stable and the other unstable. For χ further from 90, the two new fi xed points are unstable. This critical angle is approximately $\Delta\chi = 33$, so the region of "χ further from 90" is $|\chi - 90| > 33$.

The three stable zolutions for $\chi = 95$ are shown in Figure 12. At small r, the near-zero circulation is stable and in the direction of applied torque and is the result of the balance between this torque and frictional drag. For larger r, the counterclockwise circulation is in the direction of the torque, so naturally it is larger than the steady state circulation in the opposite direction.

As r increases, the stable clockwise and counterclockwise solutions undergo a Hopf bifurcation, as in the Rayleigh-Bénard case, but not simultaneously. The circulation contrary to the torque goes unstable at smaller r that in the other direction. The bifurcation diagram for $\chi = 95$ is shown in Figure 13.

Numerically computing the fi xed points and their eigenvalues gives us Figure 14, which shows the numbers of stable and unstable fi xed points in $\chi - r$ space. The transitions from 1 to 3 fi xed points are saddle-node bifurcations, and the conversion of a stable fi xed point to an unstable fi xed

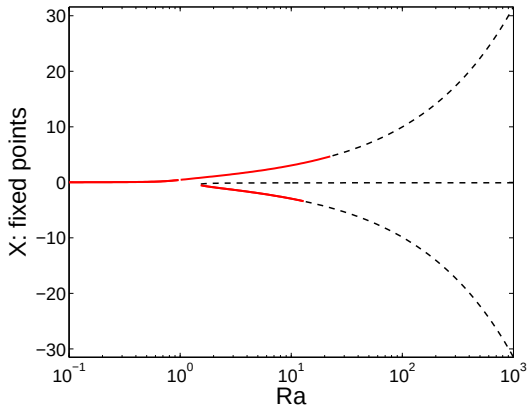


Figure 13: Bifurcation diagram for convection with $\chi = 95$. Solid lines indicate stable fi xed points; dotted lines indicate unstable fi xed points. There is an imperfect bifurcation near $r = 11$ and two Hopf bifurcations near $r = 15$ and $r = 19$.

point is a Hopf bifurcation. The curves for the locations of Hopf and saddle-node bifurcations in $\chi - r$ space may also be computed analytically from the eigenvalues of the fi xed points. The eigenvalues λ of our system are the solutions to the equation, called the characteristic polynomial,

$$det|A - \lambda I| = 0 \,, \tag{69}$$

where A is the linear stability matrix

$$A = \begin{bmatrix} -\sigma & \sigma & 0 \\ r\sin\chi - Z_0 & -1 & -X_0 \\ Y_0 & X_0 & -1 \end{bmatrix} \,, \tag{70}$$

and I is the 3×3 identity matrix. The characteristic polynomial is

$$\lambda^3 + (\sigma+2)\lambda^2 + (2\sigma + 1 + X_0^2 + \sigma(X_0 - r\sin\chi))\lambda + \sigma(1 + 2X_0^2 + X_0 - r\sin\chi) = 0 \,. \tag{71}$$

For a Hopf bifurcation, the real parts of two complex conjugate eigenvalues pass through 0. This is the same as saying that at a Hopf bifurcation, two of the eigenvalues are purely imaginary. Whether the real parts become positive or become negative determines whether or not the Hopf bifurcation is subcritical or supercritical. Then the eigenvalues at a Hopf bifurcation may be called $(\lambda_1, i\omega, -i\omega)$ giving a polynomial which must be equal to the characteristic polynomial for these to be eigenvalues of our system. Then we can solve

$$(\lambda - \lambda_1)(\lambda^2 + \omega^2) = \lambda^3 + (\sigma+2)\lambda^2 + (2\sigma + 1 + X_0^2 + \sigma(X_0 - r\sin\chi))\lambda + \sigma(1 + 2X_0^2 + X_0 - r\sin\chi) \,, \tag{72}$$

or collecting powers of λ,

$$\lambda_1 = -\sigma - 2 \,, \tag{73}$$

$$\omega^2 = 2\sigma + 1 + X_0^2 + \sigma(X_0 - r\sin\chi) \,, \tag{74}$$

$$\lambda_1\omega^2 = -\sigma(1 + 2X_0^2 + X_0 - r\sin\chi) \,. \tag{75}$$

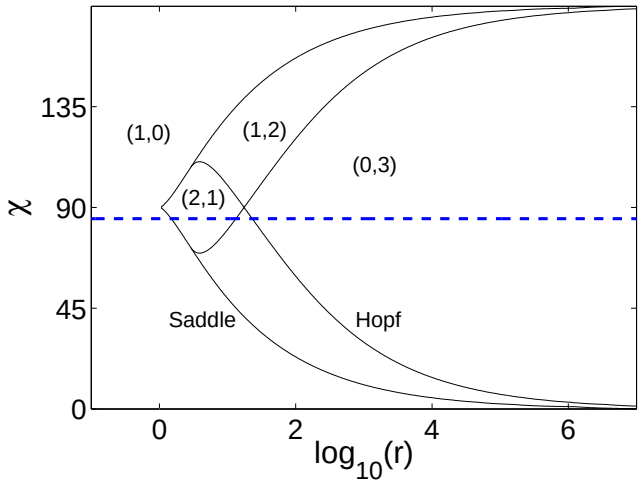


Figure 14: Bifurcation diagram in $\chi - r$ space where (*s*,*u*) indicates the numbers of stable and unstable fixed points in each region, solid lines indicate bifurcations and the dotted line is the slice through *r*-space of $\chi = 95$ for which the bifurcation diagram is shown in Figure 13.

This system of equations may then be solved for $X_0(\chi, r)$ which is the equation of the Hopf bifurcation in $\chi - r$ space. We must also require that $\omega^2 < 0$ since these are two purely imaginary eigenvalues.

Similarly, at a saddle-node bifurcation, one eigenvalue passes through zero giving the polynomial

$$\lambda(\lambda-\lambda_1)(\lambda-\lambda_2) = \lambda^3 + (\sigma+2)\lambda^2 + (2\sigma+1+X_0^2+\sigma(X_0 - r\sin\chi))\lambda + \sigma(1+2X_0^2+X_0 - r\sin\chi) \,, \tag{76}$$

so the last term must be zero, giving

$$-\sigma(1+2X_0^2+X_0 - r\sin\chi) = 0 \,, \tag{77}$$

which may be solves for $X_0(\chi, r)$, the equation of the saddle-node bifurcation. These curves are plotted in Figure 14 and precisely define the transitions between numbers of stable and unstable fixed points.

5.3 Time dependent solutions ($\chi = 45$)

Past both Hopf bifurcations, the system has gone completely unstable, so we use an explicit Runge-Kutta routine in MATLAB to numerically integrate solutions to the system. For the Rayleigh-Bénard case, which is also the Lorenz case, the system quickly becomes chaotic past the Hopf bifurcations, though much further in *r*-space reverts to periodic solutions. For $\chi = 45$, we find that the solutions are initially stable limit cycles and, at higher *r*, develop period doubling and quadrupling, and eventually chaos-like behavior as shown in Figure 15. This is indicative of a supercritical Hopf bifurcation. (In exploring the parameter space near the second Hopf bifurcation for variable χ, it appears as though near $\chi = 90$, the Hopf bifurcation is subcritical and chaos-like behavior develops

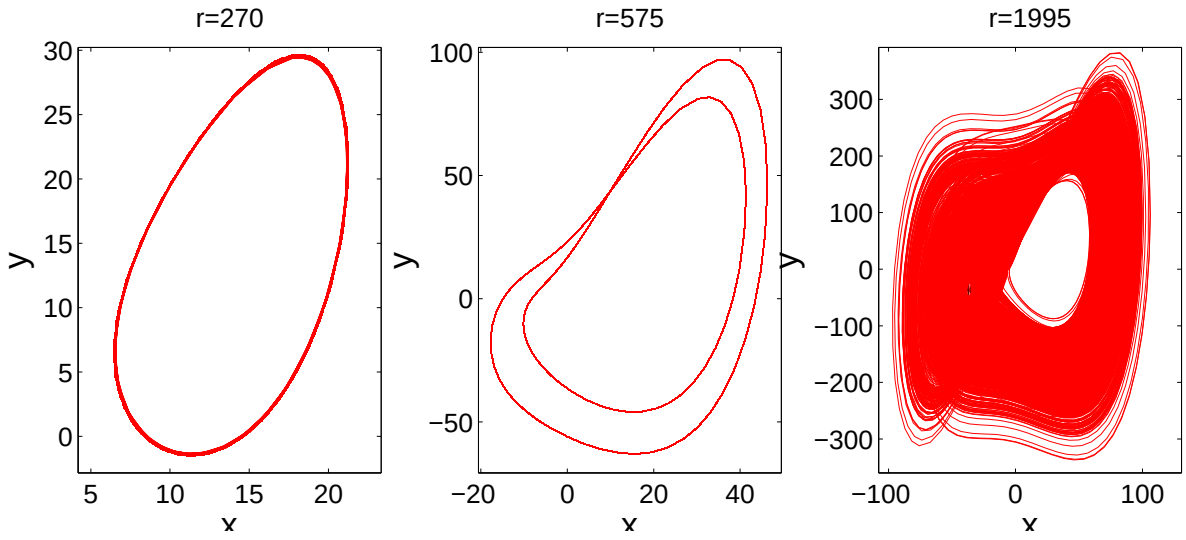


Figure 15: Phase plots of X and Y after a long time, so the initial transient is discounted. (a) For $r = 270$, (b) $r = 575$ and (c) $r = 1995$. These three plots show a stable limit cycle, period doubling and chaotic-like behavior all from the fluid loop for $\chi = 45$.

quickly in r space. Further from $\chi = 90$, by about $\chi = 75$, stable limit cycles are observed and we can conclude that the Hopf bifurcations are supercritical.) Given this sort of behavior, we can no longer explicitly calculate the circulation, dissipation and horizontal heat transport in the system.

5.4 Upper Bound on Circulation and Dissipation ($\chi > 0$)

For the case of heating below cooling ($\chi > 0$), the system exhibits limit cycles, period doubling and chaos-like behavior. However, the long-time averages of circulation, dissipation, and horizontal heat transport can be upper bounded as functions of the Rayleigh number and angle of heating χ. In our nondimensional setup given in §2.3, X is the angular velocity of the fluid. Since we are looking for an upper bound on the magnitude of circulation in the long-time averages of the system, we will look for bounds on X^2. Following the methods of Pétrélis and Pétrélis (11), we find it convenient to define

$$E \equiv \frac{1}{2}(X^2 + Y^2 + Z^2) \tag{78}$$

which has no physical meaning in our system. We also define

$$P \equiv (\sigma + r\sin\chi)XY \ , \tag{79}$$

$$D \equiv \sigma X^2 + \left(Y - \frac{r\cos\chi}{2}\right)^2 + Z^2 - \frac{r^2\cos^2\chi}{4} \ , \tag{80}$$

for which $\dot{E} = P - D$. Then if we can show that $\overline{\dot{E}} = 0$, then $\overline{P} = \overline{D}$ and from the long-time average of equation (29), we know that $\overline{X^2} = \overline{XY}$, a factor in P. So if we can find an upper bound on $\overline{D}$, we will have an upper bound on circulation.

To show that $\overline{\dot{E}} = 0$, we must simply show that E is bounded. For this we follow the methods of Doering and Gibbon (12) and construct a functional L which is decreasing for all (X,Y,Z) outside a finite ellipsoid, which means that all trajectories eventually exist inside this ellipsoid. Defining

$$L \equiv \frac{1}{\sigma}(X + r\cos\chi)^2 + Y^2 + (Z - r\sin\chi - 1)^2 \,, \tag{81}$$

its time derivative is

$$\begin{aligned}\frac{1}{2}\dot{L} &= -X^2 - r\cos\chi X - Y^2 - Z^2 + (r\sin\chi + 1)Z \,, \\ &= -\frac{1}{2}(X + r\cos\chi)^2 - Y^2 - \frac{1}{2}(Z - r\sin\chi - 1)^2 - \frac{1}{2}X^2 - \frac{1}{2}Z^2 + \frac{1}{2}r^2\cos^2\chi + \frac{1}{2}(r\sin\chi + 1)^2 \,.\end{aligned} \tag{82}$$

Then the time derivative of L is bounded by a multiple of L and a constant term,

$$\frac{1}{2}\dot{L} \le -\frac{1}{2}L + \frac{1}{2}(r^2\cos^2\chi + (r\sin\chi + 1)^2) \,, \tag{83}$$

$$\dot{L} \le -L + 2r\sin\chi + 2 \,. \tag{84}$$

Thus outside of the ellipsoid

$$L = 2r\sin\chi + 2 \,, \tag{85}$$

the time derivative of L is negative. Thus all trajectories of X, Y and Z eventually enter this ellipsoid. Then we have shown that E is bounded and we can look for a bound on circulation,

$$\overline{X^2} = \frac{\overline{P}}{\sigma + r\sin\chi} = \frac{\overline{D}}{\sigma + r\sin\chi} \le \frac{D_0}{\sigma + r\sin\chi} \,. \tag{86}$$

The following argument will be to find a small D_0.

Using the background method of Doering and Constantin (13), we can write every solution as the sum of a constant background field and a fluctuating part in time:

$$\mathbf{X}(t) = \mathbf{X_b} + \mathbf{x}(t). \tag{87}$$

Substituting this into equation (80) and taking the long time average, we can write

$$\overline{D} = 2(F(\mathbf{X_b}) - \overline{Q(\mathbf{X_b}, \mathbf{x})}) \,, \tag{88}$$

where

$$F(\mathbf{X_b}) = \frac{1}{2}(\sigma X_b^2 + (Y_b - r\cos\chi/2)^2 + Z_b^2) \,, \tag{89}$$

$$Q(\mathbf{X_b}, \mathbf{x}) = A\hat{x}^2 + \frac{1}{2}\hat{y}^2 + \frac{1}{2}\hat{z}^2 - \frac{1}{2}X_b^2(r\sin\chi - Z_b)^2 - \frac{1}{2}X_b^2Y_b^2 - \frac{L^2}{4A} \,, \tag{90}$$

In equation (90) terms have been gathered as

$$\begin{aligned}A &= \frac{\sigma}{2} - \frac{(Z_b - r\sin\chi - \sigma)^2}{2} - \frac{Y_b^2}{2} \,, \\ L &= X_bY_b^2 + (X_br\sin\chi - X_bZ_b + \frac{1}{2}r\cos\chi)(r\sin\chi + \sigma - Z_b) + \sigma Y_b \,,\end{aligned}$$

and Q has been diagonalized in the fluctuating terms with

$$\begin{aligned}
\hat{x} &= x - \frac{L}{2A} \ , \\
\hat{y} &= y + (Z_b - r\sin\chi - \sigma)x + X_b(Z_b - r\sin\chi) - \frac{1}{2}r\cos\chi \ , \\
\hat{z} &= z - Y_b(X_b + x) \ .
\end{aligned}$$

In order to bound D_0 from above, we must be able to bound Q from below which means that we need $A \geq 0$. For $A \geq 0$, Q is a upward facing paraboloid in (X,Y,Z)-space so the minimum of Q is the apex of the paraboloid. If $A = 0$, then we must have $L = 0$ to avoid singularities in the solution. Since $Q(\mathbf{X_b}, \hat{\mathbf{x}} = 0) \leq Q(\mathbf{X_b}, \hat{\mathbf{x}})$, then

$$\overline{D} \leq D_0 = 2(F(\mathbf{X_b}) - Q(\mathbf{X_b}, \hat{\mathbf{x}} = 0)) \ . \tag{91}$$

Now we would like to pick an $\mathbf{X_b}$ which minimizes D_0. One such solution is

$$\mathbf{X_b} = \quad \frac{r\cos\chi}{2(\sigma - \sqrt{\sigma})}, 0, r\sin\chi + \sigma - \sqrt{\sigma} \tag{92}$$

which gives a relatively tight bound on circulation,

$$\overline{X^2} \leq \frac{D_0}{\sigma + r\sin\chi} = \frac{\frac{1}{4}\frac{\sigma}{(\sigma-\sqrt{\sigma})^2}r^2\cos^2\chi + (r\sin\chi + \sigma - \sqrt{\sigma})^2}{r\sin\chi + \sigma} \ . \tag{93}$$

Using this choice of $\mathbf{X_b}$ gives the bounds for heating below cooling of

$$\overline{X^2} \leq O(r) \ , \tag{94}$$

$$\bar{\tilde{\varepsilon}} \leq O(\sigma r) \ . \tag{95}$$

These bounds, as well as numerically integrated solutions to the equations are shown in Figures 16. The bounds are quite good for large Rayleigh number $r \gg \sigma - \sqrt{\sigma}$. Solutions for $\chi > 0$ have a much more vigorous circulation than for $\chi \leq 0$, but we are still able to bound the circulation and dissipation in the system.

Horizontal heat transport may also be bounded by following the same procedure as above, but letting

$$P' = (\sigma + r\sin\chi)XY + r\cos\chi Y \ , \tag{96}$$

$$D' = \sigma X^2 + Y^2 + Z^2 \ . \tag{97}$$

Finding a bound for $\overline{D'}$ will give a bound for $\overline{P'}$ which may then be compared with the bound for $\overline{P}$ to give a bound for $\overline{Y}$. Then, noting that

$$\overline{\widehat{H_T}} = \overline{\sigma(r\sin\chi X - XZ)} = \overline{\sigma Y} \ , \tag{98}$$

a bound for $\overline{\widehat{H_T}}$ can be found.

The dimensional form of the bounds for circulation and dissipation are shown in Figure 17. Lines are the bounds for different angles and type of boundary conditions, while the stars in corresponding colors are the values computed from integrating the equations. For small γ, equivalently

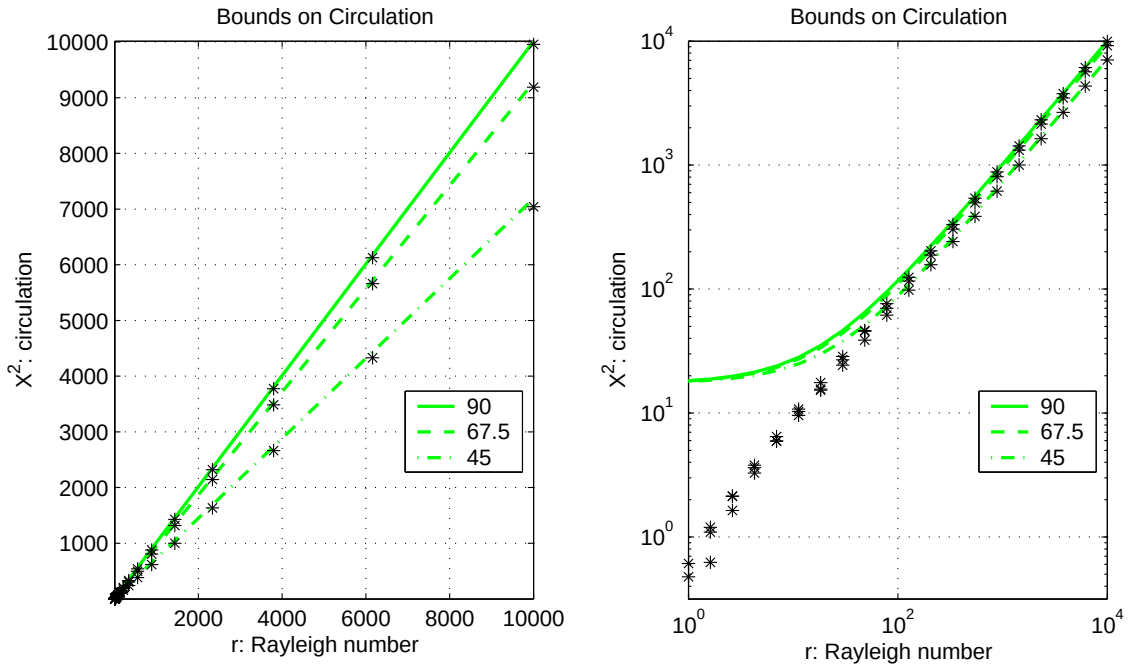


Figure 16: Heating below cooling: Linear and log plots of bounds on circulation for different angles of heating $\chi > 0$. The lines are the bounds given in equation (5.4) for $\chi = 90$, 67.5, and 45, and the stars are estimates of the quantities determined by numerically integrating the equations.

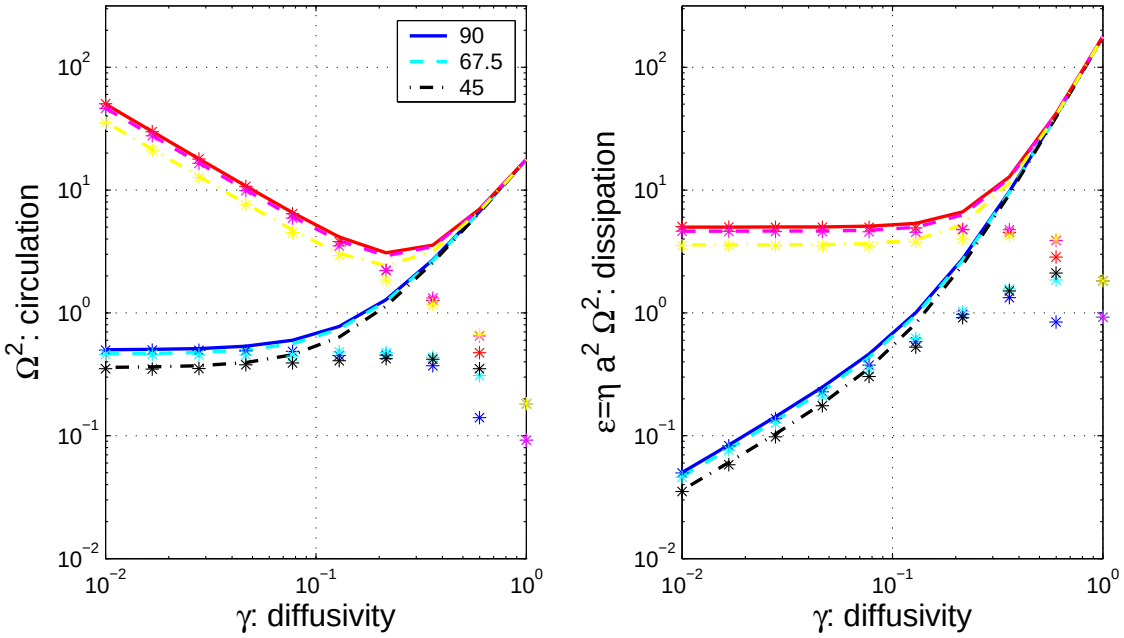


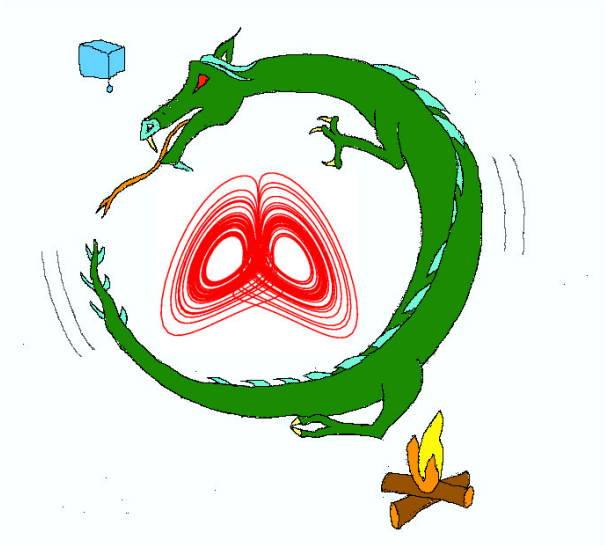
Figure 17: Heating below cooling: Log plots of dimensional bounds on (a) circulation and (b) dissipation, for different boundary conditions (fixed temperature flux in red tones, relaxation temperature in blue tones) and different angles of heating $\chi > 0$. $\chi = 90$ is a solid line, $\chi = 67.5$ is dashed, and $\chi = 45$ is dash-dot. The stars are computed for each χ by numerically integrating the equations.

large Rayleigh number, the bounds are quite good, while for larger γ the bounds are loose. In the nondiffusive limit, the circulation does not disappear and in the case of fixed flux boundary conditions, the circulation increases. The dissipation, in the thermal relaxation case, does disappear, but remains constant in the fixed flux boundary conditions. In either case, the circulation is much more vigorous than the case of horizontal convection.

6 Conclusions

Though the formulation of convection in a 1-D fluid loop is very simple, the possible solutions are diverse and complex. In the oceanic case of horizontal convection, or heating and cooling at the same level, the validity of Sandström's conjecture depends on the nature of the boundary conditions. In particular, using temperature relaxation conditions as in Paparella and Young, and most ocean GCMs, Sandström's conjecture holds: as $\kappa \longrightarrow 0$, circulation vanishes. On the other hand, using fixed buoyancy conditions, even in the limit as diffusivity vanishes, there is a substantial circulation resulting from the differential heating. If Sandström's conjecture were based on sound thermodynamic principles, one would expect it to hold regardless of the nature of the boundary conditions used.

It may be of interest to note that in fact, a variant of Sandström's theorem does hold under both thermal relaxation and fixed flux boundary conditions. In the fluid loop, the location that a parcel experiences heating (at what χ the change in potential energy following a parcel is largest) is not



the same as the location of maximally applied heating. If one considers horizontally "experienced" heating (from the point-of-view of a parcel of water), then as $\kappa \longrightarrow 0$, circulation also vanishes. This is not of much practical use in the ocean, however.

In the atmospheric analog of heating below cooling, there exist time dependent solutions with a much more vigorous circulation, though it is still possible to bound the circulation and dissipation in the system as a function of the Rayleigh number. The circulation clearly does not vanish because there is a source of potential energy, the only source which, according to Sandström's conjecture, in the absence of mixing may drive a circulation.

The main point to be taken from all this is that horizontal convection forced with buoyancy conditions (equivalently temperature and salinity) can still drive a finite circulation in the absence of diffusivity, given fixed flux boundary conditions. To motivate the relevance of this statement for the oceans, we note that though sensible heat transfer from the atmosphere to the ocean behaves as a relaxation temperature condition, latent heat and salinity forcing through precipitation and evaporation behave like fixed flux conditions. In fact, latent heating of the ocean is more than twice the magnitude of sensible heating of the ocean, so we would expect fixed buoyancy flux boundary conditions to be more appropriate.

In a fluid loop, a purely thermohaline circulation can exist in the absence of mechanical energy input from wind stress or tidal dissipation. It is not clear that a purely thermohaline circulation cannot exist in the absence of external mechanical energy forcing.

7 Acknowledgments

The author would first like to thank Prof. Raffaele Ferrari of MIT for suggesting the problem and for generous assistance, explanations and suggestions at every step of the way, as well as the Geophysical Fluid Dynamics program at Woods Hole, MA for funding and providing an environment so conducive to study. In addition, many helpful comments and suggestions came from Profs C. Doering, R. Kerswell, N. Lebowitz, and E. Spiegel.

References

[1] W. Young, personal communication (unpublished).

[2] W. H. Munk and C. Wunsch, "Abyssal Recipes II: Energetics of tidal and wind mixing," Deep Sea Res. I **45**, 1977 (1998).

[3] R. X. Huang, "Mixing and energetics of the oceanic thermohaline circulation," J. Phys. Ocean. **29**, 727 (1999).

[4] A. Defant, *Physical Oceanography I* (Pergamon Press, Oxford, 1961), p. 729.

[5] F. Paparella and W. R. Young, "Horizontal convection is non-turbulent," J. Fluid Mech. **466**, 205 (2002).

[6] W. V. R. Malkus, "Non-periodic convection at high and low prandtl number," Mem. Soc. R. Sci. Liège 6^e **IV**, 125 (1972).

[7] P. Welander, "On the oscillatory instability of a differentially heated fluid loop," J. Fluid Mech. **29**, 17 (1967).

[8] P. Welander, in *Large-scale transport processes in oceans and atmosphere*, edited by J. Willebrand and D. L. T. Anderson (NATO ASI Series, Reidel, 1986), pp. 163–200.

[9] D. J. Tritton, *Physical fluid dynamics* (Oxford University Press, Oxford, 1988).

[10] P. G. Drazin, *Nonlinear Systems* (Cambridge University Press, Cambridge, 1992).

[11] F. Pétrélis and N. Pétrélis, "Bounds on the dissipation in the lorenz system," Physics Letters A **326**, 85 (2004).

[12] C. R. Doering and J. D. Gibbon, *Applied Analysis of the Navier-Stokes Equations* (Cambridge University Press, Cambridge, 1995).

[13] C. R. Doering and P. Constantin, "Variational bounds in dissipative systems," Physica D **82**, 221 (1995).

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO. WHOI-2005-08	2.	3. Recipient's Accession No.

4. Title and Subtitle
2004 Program of Study: Tides

5. Report Date
July 2005

6.

7. Author(s) Neil J. Balmforth and Stefan Llewellyn-Smith Co-Directors;
Myrl Hendershott and Christopher Garrett, Principal Lecturers

8. Performing Organization Rept. No.
WHOI-2005-08

9. Performing Organization Name and Address

Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543

10. Project/Task/Work Unit No.

11. Contract(C) or Grant(G) No.
(C) N00014-04-1-0157
(G) OCE-0325296

12. Sponsoring Organization Name and Address

Office of Naval Research
National Science Foundation

13. Type of Report & Period Covered
Technical Report

14.

15. Supplementary Notes
This report should be cited as: Woods Hole Oceanog. Inst. Tech. Rept., WHOI-2005-08.

16. Abstract (Limit: 200 words)

The summer of 2004 saw the GFD program tackle "Tides." Myrl Hendershott (Scripps Institution of Oceanography) gave a fabulous introduction to the subject in the first week of the course, laying the foundations from astronomy and classical geophysical fluid dynamics. In the second week, Chris Garrett (University of Victoria) admirably followed up with recent developments on the subject, including the recent observations from satellite altimetry, their implications to mixing and circulation, and even a memorable lecture on the noble theme of how we might solve the world's energy crisis. The principal lectures proved unusually popular this summer, and the seminar room at Walsh often overflowed in the first two weeks.

Following on from the lectures, the seminar schedule of the summer covered in greater detail the oceanographic issues with which researchers are actively grappling. We also heard about related problems regarding atmospheric, planetary and stellar tides, together with the usual mix of topics on GFD in general.

17. Document Analysis **a. Descriptors**
tides
internal waves
ocean tides

b. Identifiers/Open-Ended Terms

c. COSATI Field/Group

18. Availability Statement
Approved for public release; distribution unlimited.

19. Security Class (This Report)
UNCLASSIFIED

20. Security Class (This Page)

21. No. of Pages
326

22. Price

(See ANSI-Z39.18) *See Instructions on Reverse*

OPTIONAL FORM 272 (4-77)
(Formerly NTIS-35)
Department of Commerce

DOCUMENT LIBRARY

Distribution List for Technical Report Exchange – *July 1998*

University of California, San Diego
SIO Library 0175C
9500 Gilman Drive
La Jolla, CA 92093-0175

Hancock Library of Biology & Oceanography
Alan Hancock Laboratory
University of Southern California
University Park
Los Angeles, CA 90089-0371

Gifts & Exchanges
Library
Bedford Institute of Oceanography
P.O. Box 1006
Dartmouth, NS, B2Y 4A2, CANADA

NOAA/EDIS Miami Library Center
4301 Rickenbacker Causeway
Miami, FL 33149

Research Library
U.S. Army Corps of Engineers
Waterways Experiment Station
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Marine Resources Information Center
Building E38-320
MIT
Cambridge, MA 02139

Library
Lamont-Doherty Geological Observatory
Columbia University
Palisades, NY 10964

Library
Serials Department
Oregon State University
Corvallis, OR 97331

Pell Marine Science Library
University of Rhode Island
Narragansett Bay Campus
Narragansett, RI 02882

Fisheries-Oceanography Library
151 Oceanography Teaching Bldg.
University of Washington
Seattle, WA 98195

Library
R.S.M.A.S.
University of Miami
4600 Rickenbacker Causeway
Miami, FL 33149

Maury Oceanographic Library
Naval Oceanographic Office
Building 1003 South
1002 Balch Blvd.
Stennis Space Center, MS, 39522-5001

Library
Institute of Ocean Sciences
P.O. Box 6000
Sidney, B.C. V8L 4B2
CANADA

National Oceanographic Library
Southampton Oceanography Centre
European Way
Southampton SO14 3ZH
UK

The Librarian
CSIRO Marine Laboratories
G.P.O. Box 1538
Hobart, Tasmania
AUSTRALIA 7001

Library
Proudman Oceanographic Laboratory
Bidston Observatory
Birkenhead
Merseyside L43 7 RA
UNITED KINGDOM

IFREMER
Bibliothèque La Pérouse
Centre de Documentation sur la Mer
15 rue Dumont d'Urville
Technopôle Brest-Iroise
BP 70 — 29280 Plouzané
FRANCE